

Db 1 mbwikklltafcivivqvhssgsfelrlklyfsndhgrdnegrcssgesdgatgkclgs 60
Qy 1 MWWIKLLTAFAICTVIVVQVHSSGSFELRLKYFSNDHGRDNEGRCCSGSDGATGKCLGS 60
Db 61 ektrfrvclkhqatidttstqctgydvitpilgensvntdgrfqnkgfntpiqpfpsf 120
Qy 61 CXTRFRVCLKHQATIDTTSTQCTGYDVITPILGENSVNLTDQRFQNGKFTNPQPFPSF 120
Db 121 swpftsliveawhdtntnsgnartnklllqrlvqvqvlevsswtknksesqytsleydf 180
Qy 121 SWPFTSLIVEAWHDTNNSGNARTNKLQRLVQVQVLEVSSWTKNKSQYTSLEYDF 180
Db 181 rvtcdlnvygscakfcrprddsfghstcsetgeiicltgwgqdychpkcagcehgbc 240
Qy 181 RVTCDLNYGSCAKFCRPRDVSFGHSTCSETGEIICLTGWQGDYCHPKCAGCEHGBC 240
Db 241 dkrcnvcqglwkgalcnevclepncihgctcnkpwcticnecwgglycncqldlnyctnhrp 300
Qy 241 DKRCNVCQGLWKGALCNEVCLEPNCHGTCNKPWCTICNEGWGGLYCNDLNYCTNHRP 300
Db 301 cknggtcftcgegllytckcapysgddceneiyscdadvnncqnggtctidephtktykc 360
Qy 301 CKNGGTCFTCGEGLYTCCKAPYSGDDCENEIYSCDADVNPQNGGTCIDEPHTKTYKC 360
Db 361 hcangwsgkmeekvltcsdkpckhgcicrnvrpqlgskqggycccpigysgncdlql 420
Qy 361 HCANGWSGKMEKVLTCSDKPCKHGCICRNVRPQLGSKQGGYCCCPIGYSGNCDLQLD 420
Db 421 ncsnpncingscqpsgkicpagfsgrtcetnidclghqcenggtctidmwnqyrcqv 480
Qy 421 NCSNPNCINGSCQPSGKICPAGFSGRTCETNIDCLGHQCENGGTCIDMWNQYRCQV 480
Db 481 pfghthcskvdclirpcanggtcclnndyqctcragftgkdcsvdidesgpcnha 540
Qy 481 PGFHTHCSKVDCLIRPCANGGTCCLNNDYQCTCRAGFTGKDCSVDIDESGPCNHN 540
Db 541 gqtcnmrvnsfecvcangfrgkqcdesvdfdhqygatqaradgltinaqvvlav 600
Qy 541 GQTCNMRVNSFECVCANGFRGKQCDSESDVDFDHQYGATQARADGLTINAQVVLAV 600
Db 601 fsvampvlavaiacvcmkrkrkraekddaearkqnegnavattmhngsgvgvalasa 660
Qy 601 FSVAMPLVAVIAACVFCMKRKRKRAEKDDAEARKQNEGNAVATMHHNGSGVGVALASA 660
Db 661 slggtksnslitfdggnpnlkntwdksvnnicasaasaaaaadeclmyggyvasv 720
Qy 661 SLGGTKSNSLITFDGGNPNIKNTWDKSVNNICASAAAAAADADECLMYGGYVASV 720
Db 721 adnnnansvfcvaplqrakskqlntdptlmhrgspagsakgagsggpgaaegkrisvl 780
Qy 721 ADNNNANSDFCVAPLQRAKSKQLNTDPTLMHRGSPAGSAGAGTAAQOQSVVCGTPTH 780
Db 781 gqgyscqrwpslaaagvagacssqlmaasaaagsagtaqqrsvvvgctphm 833
Qy 781 GQGYSCQRWPSLAAGVAGACSSQLMAASAAAGSAGTAAQOQSVVCGTPTH 833

RESULT 2
ID W11719 standard; Protein: 727 AA.
AC W11719;
DC 28-APR-1997 (first entry)
DE C-Delta-1 polypeptide.
KW C-Delta-1; cell proliferation; nervous system disorder;
KW tissue regeneration; Notch; cervix cancer; breast cancer;
KW lung cancer; colon cancer; melanoma; seminoma;
KW neurogenesis; therapy.
OS Gallus sp.
FH Key Location/Qualifiers
FT domain 184..228
FT domain /label= DSL
FT domain 229..261
FT domain /label= EGF1
FT domain 262..292

FT domain /label= EGF2
FT 293..332
FT /label= EGF3
FT 333..370
FT /label= EGF4
FT 371..409
FT /label= EGF5
FT 410..447
FT /label= EGF6
FT 448..485
FT /label= EGF7
FT 486..523
FT /label= EGF8
FT 524..534
FT /label= EGF9
FT 535..579
FT /label= TM
FT /note= *transmembrane domain*
PN W09701571-A1.
PD 16-JAN-1997.
PF 28-JUN-1996; U11178.
PR 28-JUN-1995; US-000589.
PA (IMCR) IMPERIAL CANCER RES TECHNOLOGY.
PA (UYVA) UNIV YALE.
PI Artavanis-Tsakonas S, Gray GE, Henrique D, Ish-Horowitz D,
DR WPI; 97-100159/09.
DR N-PSDB; T58897.
PT New vertebrate Delta protein, DNA and antibodies - for treating and
PT preventing cancer, nervous system disorders and for tissue
PT regeneration
PS Disclosure; Fig 2: 135pp; English.
CC C-delta-1 polypeptide (W11719) is the chick homologue of Drosophila
CC Delta, a protein that binds to Notch protein. Expression of
CC C-Delta-1 correlates with onset of neurogenesis. The C-delta-1
CC amino acid sequence was deduced from a cDNA clone (T58897) obtd.
CC from chick stage 4-6 embryos. An alternatively spliced variant
CC (W00876) was also isolated, and mouse (W11720) and human (W11721-
CC 38) Delta-1 polypeptides have been identified. Delta-1 proteins
CC can be used to treat or prevent disorders characterised by
CC increased Notch activity, such as cervical, breast, lung or colon
CC cancer, melanoma or seminoma, and nervous system disorders or to
CC promote tissue regeneration and repair.
SQ Sequence 727 AA;
Query Match 31.0%; Score 1910; DB 21; Length 727;
Best Local Similarity 50.1%; Pred. No. 9.95e-151;
Matches 263; Conservative 78; Mismatches 168; Indels 16; Gaps 15;
Db 20 qvdsgvfeiklqefvnnkgl1snrccrggpggagqgqcdctkffrvclkhyaqavsp 79
Qy 19 QVHSSGSFELRLKYFSNDHGRDNEGRCCSGSDGATGKCLGSKCTFRVCLKHQATIDT 78
Db 80 epbctgysaitpvlqansfsvpdgagadpafsnpirfpgftwpqtfsliealhtdsp 139
Qy 79 TSQCTTGVDITPILGENSVNLTDQRFQNGKFTNPQPFPSWPGFTSLIVEAWH-DTN 137
Db 140 dlttenperlisrlatqrhlavgeewsqdlhssgrtdlkyysrfvdehyygecsvfc 199
Qy 138 NSGNARTNKLQRLVQVQVLEVSSWTKNKSQYTSLEYDFRVTCDLNYGSGCAKFC 197
Db 200 rpdrdrfghfcgckvncpnpwkgqyctepiclpqcdqhgfdckpgeckervvgqr 259
Qy 198 RPRDSDFGHSTCSETGEIICLTGWQGDYCHPKCAGKC-E-HGHCDDKPNQCVQLGKGA 255
Db 260 ydcdecirypcglhtcqqpwqncqcgqgglfncqndlnycthhkpcnkatcntggsy 319
Qy 256 LCNECVLEPNCHGTCNKPWCTICNEGWGGLYCNDLNYCTNHRPCKNGGTCNTGGLY 315
Db 320 tscrcrpgytsceieinecdad--npcknggscdt-le-ns-ysctcpgpfygnkelsa 374
Qy 316 TCKCAPGYSDDCENEIYSCDADVNPQNGGTCIDEPHTKTYKCHCANGSGKMECEKV 375

	Best Local Similarity 49.1%:	Pred. No. 1,44e-144:
Matches 256; Conservative 80; Mismatches 164; Indels 21; Gaps 17;		
Db	1 sgvfeklqgefwnkkgyllgnncrrga-gppp-c--actffvclkygasvspepc 56	
Oy	23 SGEFELRLKFTYSNDRNNEGRCSSGESDQATGKLGCKTRFVCLKHQATIDTTSQC 82	
Db	57 tygasvtrvlyavdftslpddgg-adsatsnrlrfpfgtwpgttlliealhdspddla 115	
Oy	83 TYGDVPTILIGENSNTLDAQRFQNGKGTNTIOEPFSFSWGTSTSLVEAMH-DTNSGN 141	
Db	116 temperllsrlatgrhltygeewsqdlhsgrtdlksyrfvcdelhygegcsvfcrprd 175	
Oy	142 ARTNKLRLRLVQQVLEVSSEMKTNKSESOYTSLEIDFRYCDLNTYSGCAFCRPRD 201	
Db	176 daqfhtgergergekvcnpwykqpyctepiclpqcdqhgfdckpgeckcrvgygrycde 235	
Oy	202 DSFGHSTCESEGEIICLTLMGQGDYCHLIPKCAKGC-E-HGCDKRPNOCVCOLGMKGAICNE 259	
Db	236 clirpygcrlhgcqgppwgncgeqgyglfcngdlycthhkpkkgacttltqsgyscsc 295	
Oy	260 CVLEPNCIHHGICNKPWCTICNEGWGLVCNDLNTNHRKPKNGKGCFTNGBSLYTCKC 319	
Db	296 rpytygacelgidecdp--spckngsgacta-le-n-s-ysectcpgyfygacelsamtca 350	
Oy	320 APGYSDDCEMEITSCQADVAPNQCNGGTCIDEPHTKTYGKHCANAGMSGKMKCEKVLTS 379	
Db	351 dqpccfnng-r-c-sd-spdp-gysacrepvgysgfncelkldyesspscnagacvdlgda 405	
Oy	380 DKPCFHOGICRNVVRPLGSKGOGYCECPIGYSGPNCIDLQDNCSPNFCINGSGCPSGK- 438	
Db	406 yllccgagfsgrrhddndvccaspcpangfcrdgvndfscetppgytygncapvrcye 465	
Oy	439 --CICPASFSTRCCTNTIDCLGHOCENGCGICIDMWNQYRCQVGFHGHCSKVDLCL 496	
Db	466 hapcnagatchergthyryccacarygspncflllpelppp 506	
Oy	497 IRPCANGTCLTNNDYQCTCRAGFTGDCSVDIDECSSGP 537	
RESULT 7		
ID	W11720 standard; Protein; 722 AA.	
AC	W11720:	
DT	28-APR-1997 (first entry)	
DE	M-Delta-1 polypeptide.	
KW	M-Delta-1: cell proliferation; nervous system disorder;	
KW	tissue regeneration; Notch; cervix cancer; breast cancer;	
KW	lung cancer; colon cancer; melanoma; seminoma;	
OS	neurogenesis; therapy.	
KS	Mus sp.	
PN	W09701571-A1.	
PD	16-JAN-1997.	
PF	28-JUN-1996; U11178.	
PR	28-JUN-1995; US-000589.	
RA	(IMCR) IMPERIAL CANCER RES TECHNOLOGY.	
PA	(UYTA) UNIV YALE.	
PI	Ardevanis-Tsakonas S, Gray GE, Henrique D, Ish-Horowicz D;	
PI	Lewis J;	
DR	WPI: 97-100159/09.	
DR	N-PSDB: T58899.	
PT	New vertebrate Delta protein, DNA and antibodies - for treating and	
PT	preventing cancer, nervous system disorders and for tissue	
PT	regeneration	
PS	Claim 4; Fig 8; 135pp; English.	
CC	M-delta-1 polypeptide (W11720) is the mouse homologue of Drosophila	
CC	Delta, a protein that binds to Notch protein. It is expressed	
CC	primarily in presomitic mesoderm, the central and peripheral	
CC	nervous systems, and kidney. Chick (W11719) and human (W11721-	
CC	38) Delta-1 polypeptides have also been identified. Delta-1	
CC	proteins can be used to treat or prevent disorders characterised by	
CC	increased Notch activity, such as cervical, breast, lung or colon	
CC	cancer, melanoma or seminoma, as well as nervous system disorders,	
CC	and to promote tissue regeneration and repair.	

Query Match	Best Local Similarity	Score	DB 21:	Length	722:
Matches 254;	Conservative 87;	Mismatches 166;	Indels 22;	Gaps	
14 alicvswssgyfelfkigefvfkkgkq1lgnrncrcrgs-gpp--c-aactfrvclkhya	68				
15 tyivvghssgsfelrlkfrfnsdghrdngrrccspsdsdatkrclsgctrrrvclkhya	74				
69 svspsectysavprvlgvdsfsipdaq1-dpafsnlrfpfqfwpqyslllealh	127				
75 tldttsctgygvnlpilgensvnltdqrofronkgftrpiqfepfsswpgfsslveamh	134				
128 tdsppdlatemperlstrltqrlhltvgaeesqdlhnsgrdrlsyryfvedehyyegc	187				
135 -DTNNSGNARNTKLLIQRLVQOVLEVSSSEWTKNKSQYLSLEYDFRYTCDDLWYVSGC	193				
188 svyfcprddafgftcvgdrgemcdpvgkgyqctdplc1p9cdqghyckpgeckcrvg	247				
194 AKFCPRDDSEHSCISSETGELICLTGWSGYCH1PKCAKCE--HGHCDEPNCQVQLG	253				
248 wgyrycdccitrypgclhctqgwgwcnqegwsg1fcdnqldnycthkhpcrnatctntg	307				
252 WKGALCNECVLEPNCINHGTCNKPWTIC1NEGWSGLYCNQDLYNTNRHPCKNGGTCEFTG	311				
308 qgsyrcscprygtancleavdec-ap-spknqasctd-1ed-s-1scctppfykykc	365				
312 BELYVLCRKAAPGSGDCENEYS1CDADVNPQNGGTCLDEPHRTGYKCHCANWSGRMC	371				
363 e1samtcdagpcfnng-f-c-sd-npd4-gytcnplp1fsgfneckmldcgsspscsa	417				
372 EKRVLTCSDKRCQHCICNNRPGLSKRGQGCPCPICISYSPNCDLQDNCSPPRC1NGG	433				
418 kvcdlgnsylicrcgafgyrycednvdcasspaangtcrdsyvdtsctcpypgytgc	477				
432 SCQPSGK---C1CPAGFSGTRCETENIDDC1LCHQCENGGTCD1MNOYRRCQVCPGFHGTIC	488				
478 seprvrccehaphcngatchqrgyqymcecaagyggnccqfl1p9pppg	526				
489 SSRVCLCIRPCANGTCLN1NDYQCTCRAGTGTGKDCSDVIDRCSSGP	537				
RESULT 8					
R38305 standard; Protein; 236 AA.					
AC R38305;					
DT 30-NOV-1993 (first entry)					
DE Sequence of a delta protein.					
KW Delta; topolythmic protein; family.					
OS Drosophila melanogaster.					
PN MO3312141-A.					
PD 24-JUN-1993.					
PF 11-DEC-1991; U09240.					
PI 11-DEC-1991; MO-U09240.					
PA (UYVA) UNIV YALE.					
PI Artavanis-tsakonas S, Fleming RJ;					
DR WPI; 93-214095/26.					
DR N-PSDB; Q43911.					
PT Purified serrate protein, nucleic acid and antibodies - used in					
PT the study and manipulation of differentiation and other					
PT physiological processes					
PS Disclosure, Page 85; 119pp; English.					
CC Delta encodes an approx. 100 kd protein (Delta denotes 'DLZ', the					
CC protein product of the predominant zygotic and maternal					
CC transcripts) that has nine EGF-like repeats within its					
CC extracellular domain. Molecular studies have lead to the suggestion					
CC that Notch and Delta constitute biochemically interacting elements					
CC of a cell communication mechanism involved in early developmental					
CC decisions.					
CC Sequence 236 AA;					
Query Match 28.9%; Score 1784; DB 7: Length 236;					
Best Local Similarity 100.0%; Pred. No. 8.50e-140;					

```
Matches 236; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1 mbwiklltalfictvivqvhsgsfelrlkyfsndhgrdnegrccsgesdgatgkclgs 60
Qy 1 MHWIKLLTAFICTFVIVQVHSSGSEFELRLKYFSNDHGRDNegrccsgesdgatgkclgs 60
Db 61 cktfrvclkhvqatidttscqtgvdvtpilgensvnltdaqrnfqngftnpgfjfsf 120
Qy 61 CKTFRVCLKHVQATIDTTSCQTYGVDVTPILGENSVNLTDQRFQNGFTNPQFPF 120
Db 121 svpgtfsllvesahdtnnsgnartnkllqrlivqgvlevsswknksesqytsleydf 180
Qy 121 SWPGTFSLLVEAHDTNNSGNARTNKLLQRLVQGVLEVSSWKNKSesQYTSLEYDF 180
Db 181 rvtcdlnyygscakrcprddsfghstcsetgeiicltgwqdychipkacgce 236
Qy 181 RVTCDLNYGSCAKRCPRDDSFHSTCSETGEIICLTGWQDYCHIPKARGCE 236

RESULT 9
ID W05835 standard; Protein; 1193 AA.
AC W05835;
DT 28-JAN-1997 (first entry)
DE Chick Serrate.
KW C-Serrate; Notch; cell differentiation; cell fate; tissue repair;
KW central nervous system; cancer; therapy; diagnosis.
OS Gallus sp.
FH Key
FT Location/Qualifiers
FT 1..1041
FT /label= Extracellular_domain
FT 1..5
FT /label= Sig_peptide
FT /note= "lacks the N-terminal portion owing to
FT truncation of the encoding cDNA clone"
FT 158..203
FT /label= DSL
FT /note= "region of homology with Drosophila Delta
FT and Serrate, predicted to mediate binding
FT with Notch"
FT 208..837
FT /label= ELR
FT /note= "epidermal growth factor-like repeat domain"
FT 208..238
FT /label= ELR1
FT 239..274
FT /label= ELR2
FT 275..313
FT /label= ELR3
FT 314..351
FT /label= ELR4
FT 352..390
FT /label= ELR5
FT 391..427
FT /label= ELR6
FT 428..464
FT /label= ELR7
FT 465..502
FT /label= ELR8
FT 503..540
FT /label= ELR9
FT 541..606
FT /label= ELR10
FT 607..644
FT /label= ELR11
FT 655..682
FT /label= ELR12
FT 683..721
FT /label= ELR13
FT 722..759
FT /label= ELR14
FT 760..797
FT /label= ELR15
FT 798..837
```

```
FT region /label= ELR16
FT 854..911
FT /label= Cysteine-rich_region
FT domain 1042..1066
FT /label= Transmembrane_domain
FT domain 1067..1193
FT /label= Intracellular_domain
PN W09627610-AL.
PD 12-SEP-1996.
PF 07-MAR-1996; U03172.
PR 07-MAR-1995; US-400159.
PA (IMCR ) IMPERIAL CANCER RES TECHNOLOGY.
PA (UYA ) UNIV YALE.
PI Artavanis-Isakonas S, Gray GE, Henrique DMP, Ish-Horowicz D;
PI Lewis JH, Mann RN, Myat AM;
DR WPI; 96-425379/42.
DR N-PSDB; T40092.
PT Vertebrate Serrate protein and related DNA - used to treat or
PT prevent malignancies characterised by increased Notch activity.
PS Disclosure; Page 112-115; 161pp; English.
CC Chicken Serrate (W05835), or C-Serrate, is a ligand for the zygotic
CC neurogenic locus Notch and is believed to play a major role in
CC determining cell fates in the central nervous system. Its amino
CC acid sequence was deduced from a cDNA clone (T40092) obtd. from an
CC optic explant cDNA library. C-Serrate is expressed in the central
CC nervous system, cranial placodes, nephric mesoderm, vascular
CC system, and limb bud mesenchyme.
SQ Sequence 1193 AA;

Query Match 27.9%; Score 1718; DB 19; Length 1193;
Best Local Similarity 43.1%; Pred. No. 4,46e-134;
Matches 239; Conservative 103; Mismatches 184; Indels 28; Gaps 19;

Db 1 qvasasgqfeilelsqvngvnlqngnccdgtrnpgdkctktrdecctyfkvclkeysrv 60
Qy 19 QVHS-SGSFELRLKYFSNDHGRDNegrccsgesdgatgkcl-GSCKTRFRVCLKHQAT 76
Db 61 tagpccsfsgkstpvlgntfnlkys-r--nne-knripvipffawprsytlivew-dy 115
Qy 77 DTTSQCTYGVITPILGENSVNLTDQRFQNGFTNPQFPFSGFTSLIVEAWHDT 136
Db 116 ndns-tnpdr-i-lekashsgminpsrqtqkltkhtgaahfeyqirvtcaehyqfgcnkf 173
Qy 137 NNSGNARTNKLLQRLVQGVLEVSSWKNKSesQYTSLEYDFRVTCDLNYGSGCAKF 196
Db 174 crprddffthhtcdqngnktclegtwqpecnkaicrgcspkbgstcvpgectqyqwg 233
Qy 197 CRPRDDSFHSTCSETGEIICLTGWQDYCHIPKARGCE--HGHCCKPQCVCQLGKWG 254
Db 234 qycdkcihpqgvhgtcielpwqclcetnwggqlcdkldnycgthppclnggtsentgpd 293
Qy 255 ALNCEVLEPNCIHGTCNKPTWTCICNMGWGLYCNQDLNCTNHRPKNGKTCFNTGEGL 314
Db 294 yqcscpegysgqncelaehaclsd--pchngscl-e--tstgfecycagwagptctdn 348
Qy 315 YTCRCPAGYSGDDCENEIYSCADVNPCQNGGTCIDPHTKTKYKHCANGSWKCEEK 374
Db 349 iddcspnpcghggtcgq-d---l-v--dgfkicppqwtgktcqldanecsgkpcvnansc 401
Qy 375 VLTCSDRKC-HQICRNVRPGLSGKGQYQCECPIGYSGPNCIDLQDNCSNPNCINGSC 433
Db 402 rnlgsyyvcditgwsghncdinidorg-qcngngscrdlnvgyrcicspgyagdhcek 460
Qy 434 QP-SGK--CICPAGFSGTRCETNIDDCGLGHQCENGGTGTCIDMVNQYRCQCPVGFHGTCS 490
Db 461 dincasnpcmngghcqdcingfqclcpagfsgnlcldidycepnpcqgaqcfnlamd 520
Qy 491 KVDLCILRCPANGGTCLNLDNDYQCTCRAGFTGDKCSDVIDECSSGPCNGGTCMNRVNS 550
Db 521 yfncncpdyegknc 534
Qy 551 FECVCANGFRGKQC 564
```

ID	RESULT	10
AC	W05833	standard: Protein; 1218 AA.
DT	28-JAN-1997	(first entry)
DE	Human Serrate-1 (HJ1).	
KW	Serrate-1; human jagged-1; HJ1; Notch; cell differentiation;	
KW	cell fate; central nervous system; cancer; tissue repair; therapy;	
KW	diagnosis; antibody.	
OS	Homo sapiens.	
FT	Key	Location/Qualifiers
FT	domain	1..1067
FT	peptide	/label= Extracellular_domain
FT	domain	14..29
FT	domain	/label= Sig_Peptide
FT	FT	185..229
FT	FT	/label= DSL
FT	FT	/note= "region of homology with Drosophila Delta
FT	FT	and Serrate, predicted to mediate binding
FT	FT	with Notch"
FT	FT	234..896
FT	FT	/label= ELR
FT	FT	/note= "epidermal growth factor-like repeat domain"
FT	FT	234..264
FT	FT	/label= ELR1
FT	FT	265..299
FT	FT	/label= ELR2
FT	FT	300..339
FT	FT	/label= ELR3
FT	FT	340..377
FT	FT	/label= ELR4
FT	FT	378..415
FT	FT	/label= ELR5
FT	FT	416..453
FT	FT	/label= ELR6
FT	FT	454..490
FT	FT	/label= ELR7
FT	FT	491..528
FT	FT	/label= ELR8
FT	FT	529..566
FT	FT	/label= ELR9
FT	FT	567..598
FT	FT	/label= Partial_ELR
FT	FT	599..632
FT	FT	/label= Partial_ELR
FT	FT	633..670
FT	FT	/label= ELR10
FT	FT	671..708
FT	FT	/label= ELR11
FT	FT	709..747
FT	FT	/label= ELR12
FT	FT	748..785
FT	FT	/label= ELR13
FT	FT	786..823
FT	FT	/label= ELR14
FT	FT	824..862
FT	FT	/label= ELR15
FT	FT	863..879
FT	FT	/label= Partial_ELR
FT	FT	880..896
FT	FT	/label= Partial_ELR
FT	FT	1068..1089
FT	FT	/label= Transmembrane_domain
FT	FT	1090..1218
FT	FT	/label= Intracellular_domain
PN	W09627610-A1.	
PD	12-SEP-1996.	
PD	07-MAR-1996.	U031172.
PR	07-MAR-1995; US-400159.	
PA	(IMCR) IMPERIAL CANCER RES TECHNOLOGY.	
PA	(UYTA) UNIV YALE.	
PI	Artavanis-Tsakonas S, Gray GE, Henrique DMP, Ish-Horowitz D;	
PI	Lewis JH, Mann RS, Myat AM;	

[illegible]

```
FT Protein /label= Signal
FT 32..1218
FT /label= Differentiation_suppression_protein
PN W09719172-A1.
PD 29-MAY-1997. J03356.
PF 15-NOV-1996; JP-311811.
PR 30-NOV-1995; JP-299611.
PR 17-NOV-1995; JP-299611.
PA (ASAH ) ASAH KASEI KOGYO KK.
PI Itoh A, Sakano S;
DR WPI; 97-298110/27.
DR N-PSDB; T70175.
DR Peptide(s) encoded by human genes delta-1 and serrate-1 - suppress
PT proliferation and differentiation of undifferentiated human blood
PT cells
PS Claim 15; Page 83-91; 114pp; Japanese.
CC The present sequence represents a polypeptide which suppresses
CC proliferation and differentiation of undifferentiated cells such
CC as neurons and blood cells. The polypeptide may be used for the
CC prevention and control of disorders involving undifferentiated
CC cells, such as leukaemia and malignant tumours, and improvement of
CC blood formation, e.g. after immunosuppression.
SQ Sequence 1218 AA;

Query Match 27.3%; Score 1680; DB 25; Length 1218;
Best Local Similarity 41.5%; Pred. No. 8,74e-131;
Matches 233; Conservative 108; Mismatches 194; Indels 29; Gaps 19;

Db 16 lllallc-alrakvqasqfellelsmqnvngelqncgccgarnpgdrktrdecddy 74
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Qy 7 LLTATICTVIVQVH-SSGSFELRLKLYFSDHGRDNEGRCCSGESDGTAKCL-GSKTR 64
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :

Db 75 fkvclkyqsrvtagpcsfsgstpvigntfnlks-rgnr---nrivlpfsfawpr 130
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Qy 65 FRVCLKHYQATIDTTSQCTYGVITPILGENSVNLTDQORFQNKGTNP IQPFPSFSGP 124
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :

Db 131 sytliveav-dsdsd-tvpdsi-iekaashgminprqwtlkqntgahfeyqirvc 187
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Qy 125 TFSLIVEAWHDNNSGNATNKLRLVQVLEVSSEWTKNSQSQTSLDYDFRVTC 184
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :

Db 188 ddyvyygfnkfcprdrdfghvacdqngkctmegwmpcncraicrggscpkhgsckl 247
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Qy 185 DLNLYGSGCAKFCRPDDSFHSTCSETGEIILCTGWQGDYCHIPKCAKGE--HGCDK 242
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :

Db 248 pdrcrcqvgwqglycdkicphpgvhgicnepwqclcetnwggqldkldnycgthqpc 307
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Qy 243 PNQVCQLGWKALNECVLEPNCIHGTCNKPTWCICNEGWWGLYCNQDLNCTNHRPCK 302
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :

Db 308 nggtcsntgpkycscpegysgpnceiaehaclsd--pchnrgsc-ke--tslgfcec 362
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Qy 303 NGGTCFNTGEGLYTKCAPYSGDDCENIYSCDADVNPQNGGTCIDEPTHKTYKCHC 362
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :

Db 363 spgwgtptcstnidcsppncshggtcqd---l-vng--fkvcppqwtgkctcidanq 415
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Qy 363 ANGWSGKWEKVEKVLTCSDRPC-HQICRNVRPGLSGKGQYQCECPIGYSGPNCIDLQDN 421
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :

Db 416 ceakpcvnaeksknlasydcldplgmwmgncdinindclg-qcndascdrlvngyrci 474
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Qy 422 CSPNPGINGSCQP-SGK--CICPAGFSGTRCETNIDDCGLGHQCEGGTCIDMVNQYRCQ 478
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :

Db 475 cpbgagdhcercidceapnclngghcneinrfclcptgfsnqlcldidycenpc 534
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Qy 479 CVPFGHGTCHSSKVDJLIRPCANGGTCNLNNDYQCTCRAGFTGDKCSVDIDECSSGPC 538
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :

Db 535 qnagqcyrnasydfckcpedegknc 560
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Qy 539 HNGGTCMNRVNSFEVCANGFRGKQC 564
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
```

```
RESULT 12
ID W18351 standard; protein; 1036 AA.
AC W18351.
DI 11-FEB-1998 (first entry)
```

```
DE Proliferation and differentiation suppression polypeptide.
KW Proliferation; differentiation; suppression; human; delta-1;
KW serrate-1; blood cell; neuron; leukaemia; malignant tumour;
KW immunosuppression.
OS Homo sapiens.
PN W09719172-A1.
PD 29-MAY-1997. J03356.
PF 15-NOV-1996; JP-311811.
PR 30-NOV-1995; JP-299611.
PR 17-NOV-1995; JP-299611.
PA (ASAH ) ASAH KASEI KOGYO KK.
PI Itoh A, Sakano S;
DR WPI; 97-298110/27.
DR Peptide(s) encoded by human genes delta-1 and serrate-1 - suppress
PT proliferation and differentiation of undifferentiated human blood
PT cells
PS Claim 5; Page 66-71; 114pp; Japanese.
CC The present sequence represents a polypeptide which suppresses
CC proliferation and differentiation of undifferentiated cells such
CC as neurons and blood cells. The polypeptide may be used for the
CC prevention and control of disorders involving undifferentiated
CC cells, such as leukaemia and malignant tumours, and improvement of
CC blood formation, e.g. after immunosuppression.
SQ Sequence 1036 AA;

Query Match 27.2%; Score 1676; DB 25; Length 1036;
Best Local Similarity 42.1%; Pred. No. 1.94e-130;
Matches 231; Conservative 104; Mismatches 187; Indels 27; Gaps 17;

Db 1 sgfefeilelsmqnvngelqncgccgarnpgdrktrdecddyfkclyqsrvtagpp 60
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Qy 23 SGSEFELRLKLYFSDHGRDNEGRCCSGESDGTAKCL-GSKTRFVRLKHYQATIDTTSQ 81
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :

Db 61 csfsgstpvigntfnlks-rgnr---nrivlpfsfawprsytliveav-dsdsd-t 114
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Qy 82 CTYGDVITPILGENSVNLTDQORFQNKGTNP IQPFPSFSGPCTSLIVEAWHDNNSGN 141
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :

Db 115 vqpdsi-iekaashgminprqwtlkqntgahfeyqirvtcdydygfnkfcprdr 173
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Qy 142 ARTNKLILRLVQVLEVSSEWTKNSQSQTSLDYDFRVTCIDNLYGSGCAKFCRPRD 201
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :

Db 174 dfghvacdqngkctmegwmpcncraicrggscpkhgscklpgdcrcqvgwqglycdk 233
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Qy 202 DSFGHSTCSETGEIILCTGWQGDYCHIPKCAKGE--HGCDKPNQVCQLGWKALCNE 259
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :

Db 234 ciphpgcvhgcicnepwqclcetnwggqldkldnycgthqpcinggtcsntgpkycsc 293
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Qy 260 CVLEPNCIHGTCNKPTWCICNEGWWGLYCNQDLNCTNHRPCKNGGTCFTBGLYTCCK 319
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :

Db 294 pegysgpnceiaehaclsd--pchnrgsc-ke--tslgfcecspgwtgctcstnidcs 348
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Qy 320 APGYSGDDCENIYSCDADVNPQNGGTCIDEPTHKTYKCHANGWSGKWEKVEKVLTC 379
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :

Db 349 pnnscshggtcqd---l-vng--fkvcppqwtgkctcidaneceakpcvnaeksknlia 401
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Qy 380 DKFC-HQICRNVRPGLSGKGQYQCECPIGYSGPNCIDLQDNCSNPNGINGSCQP-SG 437
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :

Db 402 syvdcldplgmwmgncdinindclg-qcndascdrlvngyrcicpbpgagdhcercid 460
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Qy 438 K--CICPAGFSGTRCETNIDDCGLGHQCEGGTCIDMVNQYRCQVPGFHGTCHSSKVDLC 495
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :

Db 461 aspnclngghcneinrfclcptgfsnqlcldidycenpcnqagqcyrnasydfckc 520
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Qy 496 LIRPCANGGTCNLNNDYQCTCRAGFTGDKCSVDIDECSSGPCMNRVNSFEVC 555
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :

Db 521 pedegknc 529
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Qy 556 ANGFRGKQC 564
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
```

```
RESULT 13
ID W18352 standard; protein; 1187 AA.
AC W18352;
```


DT 11-FEB-1998 (first entry)
DE Proliferation and differentiation suppression polypeptide.
KM Serrate-1; blood cell; neuron; leukaemia; malignant tumour;
KW Immunosuppression.
OS Homo sapiens.
PN MO9719172-A1.
PD 29-MAY-1997.
PF 15-NOV-1996; J03356.
PR 30-NOV-1995; JP-311811.
PR 17-NOV-1995; JP-299611.
PA (ASAH) ASAH KASEI KOGYO KK.
PI Itoh A, Sakano S;
DR WPI: 97-298110/27;
PT Peptide(s) encoded by human genes delta-1 and serrate-1 - suppress
PT proliferation and differentiation of undifferentiated human blood
PT cells
PS Claim 6; Page 71-76; 114pp; Japanese.
CC The present sequence represents a polypeptide which suppresses
CC proliferation and differentiation of undifferentiated cells such
CC as neurons and blood cells. The polypeptide may be used for the
CC prevention and control of disorders involving undifferentiated
CC cells, such as leukaemia and malignant tumours, and improvement of
CC blood formation, e.g. after immunosuppression.
SQ Sequence 1187 AA;

Query Match 27.2%; Score 1676; DB 25; Length 1187;
Best Local Similarity 42.1%; Pred. No. 1.94e-150;
Matches 231; Conservative 104; Mismatches 187; Indels 27; Gaps 17;

DB 1 sgfgelellmqnyngelngncgarnpdrctedcteyfkyclkeysvrtaggp 60
QY 23 GSGFEELKLTFTSNDHGDNDNRCCSGESDGAATGKCL-GSCXTRRVCCKHQTADITTSQ 81
DB 61 csfsgstpylgntfnlkas-rgndr--nrlvlpfsfawrsyllveaw-dsmd-t 114
QY 82 CTYGDVITPLIGENSVMLTDAQRFQNGFTNPIDPFPSFSGPFSLLVEAMHMDTNSGN 141
DB 115 vqpdsl-tkxashgsmimprsqwtlqntgvahfeyqlyvtcdyyugfscnfcpd 173
QY 142 ARTNKLILIORLLVOQVLEVSSESEMTNSESQYTSLEYDFRVTCDLNYYGSGCAKFCRPPD 201
DB 174 dffghvacdngnktcmegmmgpcnraicrgqspkhscklpgddrcgywgqlycdk 233
QY 202 DSEGHSTCEGELITCLGOWGDYCHIRPKAKGE--HGCDKPNQVCQJGKMGALCNE 259
DB 234 clphgscvhlcmepwqclsetnwqglcdklnyesthqpclngtscnqrpdkyqsc 293
QY 260 CVLEPNCIHGCKNKPWTCICNEGSGLYCNODLNVCTMHRCKNGGTGFNGELGYTCKC 319
DB 294 pegysgpnclaeahclsd--pchnrgsc-ke--tsjgfecescspgvtgptcstniddcs 348
QY 320 APGYSGDCEMEIYSCADVNPCCONGGTCIDEPHTKTKYKCHCANMGSGKMCDEKVLTC 379
DB 349 pncshgtsqgd----l-vng--fkvcprpwtgktcqlaneceakpcvnaakscknli 401
QY 380 DKPC-HGIGCKNVRPGLSGKGGYQCECPISYSGPNCQLQIDNCSPPNCINGSCQP-SG 437
DB 402 syycdcipgwmgnclndclg-qcndascrdlvngyrcfpppyagdhccerdldec 460
QY 438 K--CICRAGSFGTCENIDDCIGHQCENGSTCIDMVNVOYRCQCVPRHGHNCSSKVDLC 495
DB 461 asnpclngnqhneinfqclcptqfsgnlcqlddycepnpcqngagcynraadyfckc 520
QY 496 LIRPCANGGTCLNMDNYOCTCRAGFTGKDCSDVIDECSSPCFCHNGTCMNRVNSFEVCV 555
DB 521 pedyegknc 529
QY 556 ANGFGRKQC 564

AC W05834;
DT 28-JAN-1997 (first entry)
DE Human Serrate-2 (HJ2).
KM Serrate-2; human jagged-2; HJ2; Notch; cell differentiation;
KW cell fate; central nervous system; cancer; tissue repair; therapy;
KW diagnosis; antibody.
OS Homo sapiens.
FH Key
FT domain
FT location/Qualifiers
FT 1..912
FT /label- Extracellular_domain
FT /note- "a deletion in the encoding cDNA clone
FT results in loss of part of the Serrate-2
FT signal peptide and beginning of the DSL
FT domain
FT 26..70
FT /label- DSL
FT /note- "region of homology with Drosophila Delta
FT and Serrate, predicted to mediate binding
FT with Notch"
FT 75..735
FT /label- ELR
FT /note- "epidermal growth factor-like repeat domain"
FT region
FT 75..105
FT /label- ELR1
FT 106..140
FT /label- ELR2
FT 141..180
FT /label- ELR3
FT 181..218
FT /label- ELR4
FT 219..256
FT /label- ELR5
FT 257..294
FT /label- ELR6
FT 295..331
FT /label- ELR7
FT 332..369
FT /label- ELR8
FT 370..407
FT /label- ELR9
FT 408..435
FT /label- Partial_ELR
FT 436..469
FT /label- Partial_ELR
FT 470..507
FT /label- ELR10
FT 508..545
FT /label- ELR11
FT 546..584
FT /label- ELR12
FT 585..622
FT /label- ELR13
FT 623..660
FT /label- ELR14
FT 664..701
FT /label- ELR15
FT 702..718
FT /label- Partial_ELR
FT 719..735
FT /label- Partial_ELR
FT 913..933
FT /label- Transmembrane_domain
FT 934..1257
FT /label- Intracellular_domain
PD W05627610-A1.
PD 12-SEP-1996.
PF 07-MAR-1996; U03172.
PR 07-MAR-1995; US-400159.
PA (IMCR) IMPERIAL CANCER RES TECHNOLOGY.
PA (UYVA) UNIV YALE.
PI Artavanis-Tsakonas S, Gray GE, Henrique DMP, Ish-Horowicz D;
PI Lewis JH, Mann RS, Myat AM;
DR WPI: 96-425379/42.

CC TOPOLOGY: unknown
CC MOLECULE TYPE: Protein
SQ SEQUENCE 833 AA; 88812 MW; 3155492 CN;

Query Match 100.0%; Score 6164; DB 1; Length 833;
Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 833; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 MHWIKLLTAFICFTVIVQVHSSGSEFLRLKYFSDHGRNEGRCCSGESDGTGKGLGS 60
QY 1 MHWIKLLTAFICFTVIVQVHSSGSEFLRLKYFSDHGRNEGRCCSGESDGTGKGLGS 60
Db 61 CKTRFRVCLKHQATIDTTSQCTGYDVTPLIGNSVNLDAQRFQNGFTNPIQFPFSF 120
QY 61 CKTRFRVCLKHQATIDTTSQCTGYDVTPLIGNSVNLDAQRFQNGFTNPIQFPFSF 120
Db 121 SWPGTFSLIVEAHDHNNNSGARTKLLIQRLVQVLEVSSSEWTKNKSQYTSLEYDF 180
QY 121 SWPGTFSLIVEAHDHNNNSGARTKLLIQRLVQVLEVSSSEWTKNKSQYTSLEYDF 180
Db 181 RVTCDLNYGSGCAKFCRPRDDSGHSTSETGEIICLTGQGDYCHIPKAKCEHGHC 240
QY 181 RVTCDLNYGSGCAKFCRPRDDSGHSTSETGEIICLTGQGDYCHIPKAKCEHGHC 240
Db 241 DPNQCVQGLGKALGALNECVLEPNCIHTGTCNKPTWCICNEGWGLYCNDLNTCTNHRP 300
QY 241 DPNQCVQGLGKALGALNECVLEPNCIHTGTCNKPTWCICNEGWGLYCNDLNTCTNHRP 300
Db 301 CKNGGTCFNTGEGLYTCKAPGSGDDCENEIYSCDADVNPQNGGTCIDEPHTKTYKC 360
QY 301 CKNGGTCFNTGEGLYTCKAPGSGDDCENEIYSCDADVNPQNGGTCIDEPHTKTYKC 360
Db 361 HCANGSGKCEKVLTCSDPKHQGICRNVPRGLSGKGGYQCEPIGYSGPNCDLQLD 420
QY 361 HCANGSGKCEKVLTCSDPKHQGICRNVPRGLSGKGGYQCEPIGYSGPNCDLQLD 420
Db 421 NCSPNPGINGSGQPSKGCICPAGFSCTRCETNIDDCLGHCENGCTCIDMVNRYRCQV 480
QY 421 NCSPNPGINGSGQPSKGCICPAGFSCTRCETNIDDCLGHCENGCTCIDMVNRYRCQV 480
Db 481 PGFHGTHCSKVDLCILRPCANGGTCILNNDYOCCTRAGFTGKDCSVYIDECSSGPGCHN 540
QY 481 PGFHGTHCSKVDLCILRPCANGGTCILNNDYOCCTRAGFTGKDCSVYIDECSSGPGCHN 540
Db 541 GGTGCMNRVNSFEVCANGFRGKQCEESYDSVTFDAHOYGATTOARADGLTNAQVLIIV 600
QY 541 GGTGCMNRVNSFEVCANGFRGKQCEESYDSVTFDAHOYGATTOARADGLTNAQVLIIV 600
Db 601 FSVAMPPLVAVIAACVFCMKRRKRAQEKDDAEARKQNEQNAVATMHNGSGVGVALASA 660
QY 601 FSVAMPPLVAVIAACVFCMKRRKRAQEKDDAEARKQNEQNAVATMHNGSGVGVALASA 660
Db 661 SLGGKTCGNSGLTFDGNPNIIKNTWDSVNNICASAAAAAADADECLMYGGYVASV 720
QY 661 SLGGKTCGNSGLTFDGNPNIIKNTWDSVNNICASAAAAAADADECLMYGGYVASV 720
Db 721 ADNNANSDFCVAPLQRAKSOQLNTDPTLMHRGSPAGSSAKGASGGPGAAEKRISVL 780
QY 721 ADNNANSDFCVAPLQRAKSOQLNTDPTLMHRGSPAGSSAKGASGGPGAAEKRISVL 780
Db 781 GEGSYCSQRPWPSLAAAGVAGACSSQLMAAASAGSGAGTAQOQSVVCGTPTHM 833
QY 781 GEGSYCSQRPWPSLAAAGVAGACSSQLMAAASAGSGAGTAQOQSVVCGTPTHM 833

RESULT 2
ID US-08-264-534-3 STANDARD; PRT; 203 AA.

XX xxxxxx
XX 01-JAN-1900
XX Sequence 3, Application US/08264534.

XX Sequence 3, Application US/08264534
CC Patent No. 5648464
CC GENERAL INFORMATION:
CC APPLICANT: Artavanis-Tsakonas, Spyridon et al.
CC TITLE OF INVENTION: Human No. 5648464ch And Delta, Binding Domains
CC TITLE OF INVENTION: In Toporythmic Proteins, And Methods Based Thereon
CC NUMBER OF SEQUENCES: 34
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Pennie & Edmonds
CC STREET: 1155 Avenue of the Americas
CC CITY: New York
CC STATE: New York
CC COUNTRY: U.S.A.
CC ZIP: 10036
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA: US/08/264,534
CC APPLICATION NUMBER: US/08/264,534
CC FILING DATE:
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA: US 07/695,189
CC APPLICATION NUMBER: US 07/695,189
CC FILING DATE: 03-MAY-1991
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Misrock, S. Leslie
CC REGISTRATION NUMBER: 18,872
CC REFERENCE/DOCKET NUMBER: 7326-004
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 212 790-9090
CC TELEFAX: 212 8698864/9741
CC TELEX: 66141 PENNIE
CC INFORMATION FOR SEQ ID NO: 3:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 203 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: single
CC TOPOLOGY: unknown
CC MOLECULE TYPE: peptide
CC SEQUENCE 203 AA; 22840 MW; 218129 CN;

Query Match 24.6%; Score 1514; DB 1; Length 203;
Best Local Similarity 100.0%; Pred. No. 2.06e-113;
Matches 203; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 GSFELRLKYFSDHGRNEGRCCSGESDGTGKGLSGCKTRFRVCLKHQATIDTTSOCT 60
QY 24 GSFELRLKYFSDHGRNEGRCCSGESDGTGKGLSGCKTRFRVCLKHQATIDTTSOCT 83
Db 61 YGDVITPILGNSVNLDAQRFQNGFTNPIQFPFSFQSWPGTFSLIVEAHDHNNSGNAR 120
QY 84 YGDVITPILGNSVNLDAQRFQNGFTNPIQFPFSFQSWPGTFSLIVEAHDHNNSGNAR 143
Db 121 TNKLLIQRLVQVLEVSSSEWTKNKSQYTSLEYDFRVTCDLNYGSGCAKFCRPRDSS 180
QY 144 TNKLLIQRLVQVLEVSSSEWTKNKSQYTSLEYDFRVTCDLNYGSGCAKFCRPRDSS 203
Db 181 FGHSTCSETGEIICLTGQGDYC 203
QY 204 FGHSTCSETGEIICLTGQGDYC 226

RESULT 3
ID US-08-457-135-2 STANDARD; PRT; 383 AA.

XX xxxxxx
XX 01-JAN-1900
XX Sequence 2, Application US/08457135.

CC	PatentNo. 5580738
CC	GENERAL INFORMATION:
CC	APPLICANT: LABORDA, Jorge
CC	TITLE OF INVENTION: Delta-Like Gene Expressed In
CC	TITLE OF INVENTION: Neuroendocrine Tumors
CC	NUMBER OF SEQUENCES: 9
CC	CORRESPONDENCE ADDRESS:
CC	ADDRESSEE: Foley & Lardner
CC	STREET: 3000 K Street, N.W., Suite 500
CC	CITY: Washington, D.C.
CC	COUNTRY: USA
CC	ZIP: 20007-5109
CC	COMPUTER READABLE FORM:
CC	MEDIUM TYPE: Floppy disk
CC	COMPUTER: IBM PC compatible
CC	OPERATING SYSTEM: PC-DOS/MS-DOS
CC	SOFTWARE: PatentIn Release #1.0, Version #1.25
CC	CURRENT APPLICATION DATA:
CC	APPLICATION NUMBER: US/08/597,545
CC	FILING DATE:
CC	CLASSIFICATION: 435
CC	PRIOR APPLICATION DATA:
CC	APPLICATION NUMBER: US 07/989,537
CC	FILING DATE: 11-DEC-1992
CC	ATTORNEY/AGENT INFORMATION:
CC	NAME: BENT, Stephen A.
CC	REGISTRATION NUMBER: 29,768
CC	REFERENCE/DOCKET NUMBER: 40399/166 NIH
CC	TELECOMMUNICATION INFORMATION:
CC	TELEPHONE: (202)672-5300
CC	TELEFAX: (202)672-5399
CC	TELEX: 904136
CC	INFORMATION FOR SEQ ID NO: 2:
CC	SEQUENCE CHARACTERISTICS:
CC	LENGTH: 383 amino acids
CC	TYPE: amino acid
CC	TOPOLOGY: linear
CC	ORIGINAL SOURCE:
CC	STRAIN: Human Dlx
CC	SEQUENCE 383 AA; 41201 MW; 758273 CN;
CC	SO
CC	Query Match 10.7%; Score 660; DB 1; Length 383;
CC	Best Local Similarity 40.3%; Pred. No. 1,18e-42;
CC	Matches 89; Conservative 50; Mismatches 67; Indels 15; Gaps 10.
CC	Db 33 QNCECDNDVNCRCPOGMQPLCDQCVTSPLGHLGCEPGQCTCTDGMDELCDRDYRAC 92
CC	Qy 236 EHGHCIDKRNQCVQIQGKMGALCNBCVLEPNCHGTCNKKPMTICNHEGMGLYCNQDLNVC 295
CC	Db 93 SSA-PCANNGCVSLDDGLIECSAPFGYSKGDCKRKGPCYINSGPOHSGTCYVD-EGR 150
CC	Qy 296 TNRHRCCKNGGCFNTEGSELYTCKCAPGYSGDCENEIYSCDADVNPQNSGTCTIDEPTHK 355
CC	Db 151 ASHACSLCPPEFSNGCFIVANSCTPNPCNDGVCYCTDI---G--GDPFRCPAGPIDK 203
CC	Qy 356 TGY-KCHCANWSGKMCCEKVLFTSCKRC-HQIGICRNVRPELGSKGGGTCCBPIGTSGP 413
CC	Db 204 TCSRPVINCASSPCONGSTGLQHTQVSYECICRKEFTGLTC 244
CC	Qy 414 NCDQLDNCSPNPCIINGSC-QPS--G-KCICPAGFGSTRC 450
CC	RESULT 5 STANDARD; PRT; 385 AA.
CC	ID US-08-597-545-1
CC	AC xxxxxx
CC	DT 01-JAN-1900
CC	Sequence 1, Application US/08597545.
CC	Sequence 1, Application US/08597545
CC	Patent No. 5580738
CC	CC

CC GENERAL INFORMATION:
CC APPLICANT: LABORDA, Jorge
CC TITLE OF INVENTION: Delta-Like Gene Expressed In
CC TITLE OF INVENTION: Neuroendocrine Tumors
CC NUMBER OF SEQUENCES: 9
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Foley & Lardner
CC STREET: 3000 K Street, N.W., Suite 500
CC CITY: Washington, D.C.
CC COUNTRY: USA
CC ZIP: 20007-5109
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/597,545
CC FILING DATE:
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/989,537
CC FILING DATE: 11-DEC-1992
CC ATTORNEY/AGENT INFORMATION:
CC NAME: BENT, Stephen A.
CC REGISTRATION NUMBER: 29,768
CC REFERENCE/DOCKET NUMBER: 40399/166 NIHD
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (202)672-5300
CC TELEFAX: (202)672-5399
CC TELEX: 904136
CC INFORMATION FOR SEQ ID NO: 1:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 385 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC ORIGINAL SOURCE:
CC STRAIN: Mouse Dlk
CC SEQUENCE 385 AA; 41320 MW; 772804 CN;

Query Match 10.6%; Score 654; DB 1; Length 385;
Best Local Similarity 37.0%; Pred. No. 3.63e-42;
Matches 85; Conservative 50; Mismatches 82; Indels 13; Gaps 10;
Db 26 CDPPCDPOYGFCEADNVCRHVGEGLCDKCVTAPCVNGVCKEPPQCICKDGDWGRKFC 85
QY 231 CAKGCE-H-GHCDKPNQVCQLGKWCALNCEVLEPNCIHGTCNKPWTCICNEGWGLYC 288
Db 86 EIDVRACTS-TPCANNGTCVDLEKGYECSTPGFSGKDCQHKAGPCVINGSPCQHGAC 144
QY 289 NQDLNYCTNHRPCKNGGTCFNTGEGLYTCKAPGYSGDDCENEIYSCDADVNPQNGGTC 348
Db 145 VDDEGQASHASCLCPGFGNFC--IVA-ATNSCTPNPCEN--DGVCTDIGDGRCP 199
QY 349 IDEPHTKTYKCHCANGSGKWCERKVLCTSDKPCHQGICRNVRPGL-GSKGGYQCECP 407
Db 200 AGFVDTKTSRPVSNACSGPCQNGTCLQHTQVSEFCLCKPPFMGPTCAKK 249
QY 408 IGYSGPNCDLQDNCSPNCPINGSC-QPS--G-KCICPAGFSGTRCETN 453

RESULT 6
ID US-08-457-135-1 STANDARD; PRT; 385 AA.
XX xxxxxx
AC
XX
XX
DT 01-JAN-1900
XX
DE
DE
CC Sequence 1, Application US/08457135.
CC Sequence 1, Application US/08457135
CC Patent No. 5644031
CC GENERAL INFORMATION:

CC APPLICANT: LABORDA, Jorge
CC TITLE OF INVENTION: Delta-Like Gene Expressed In
CC TITLE OF INVENTION: Neuroendocrine Tumors
CC NUMBER OF SEQUENCES: 9
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Foley & Lardner
CC STREET: 3000 K Street, N.W., Suite 500
CC CITY: Washington, D.C.
CC COUNTRY: USA
CC ZIP: 20007-5109
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/457,135
CC FILING DATE: 01-JUN-1995
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 07/989,537
CC FILING DATE: 11-DEC-1992
CC ATTORNEY/AGENT INFORMATION:
CC NAME: BENT, Stephen A.
CC REGISTRATION NUMBER: 29,768
CC REFERENCE/DOCKET NUMBER: 40399/304/NIHD
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (202)672-5300
CC TELEFAX: (202)672-5399
CC TELEX: 904136
CC INFORMATION FOR SEQ ID NO: 1:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 385 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC ORIGINAL SOURCE:
CC STRAIN: Mouse Dlk
CC SEQUENCE 385 AA; 41320 MW; 772804 CN;
Query Match 10.6%; Score 654; DB 1; Length 385;
Best Local Similarity 37.0%; Pred. No. 3.63e-42;
Matches 85; Conservative 50; Mismatches 82; Indels 13; Gaps 10;
Db 26 CDPPCDPOYGFCEADNVCRHVGEGLCDKCVTAPCVNGVCKEPPQCICKDGDWGRKFC 85
QY 231 CAKGCE-H-GHCDKPNQVCQLGKWCALNCEVLEPNCIHGTCNKPWTCICNEGWGLYC 288
Db 86 EIDVRACTS-TPCANNGTCVDLEKGYECSTPGFSGKDCQHKAGPCVINGSPCQHGAC 144
QY 289 NQDLNYCTNHRPCKNGGTCFNTGEGLYTCKAPGYSGDDCENEIYSCDADVNPQNGGTC 348
Db 145 VDDEGQASHASCLCPGFGNFC--IVA-ATNSCTPNPCEN--DGVCTDIGDGRCP 199
QY 349 IDEPHTKTYKCHCANGSGKWCERKVLCTSDKPCHQGICRNVRPGL-GSKGGYQCECP 407
Db 200 AGFVDTKTSRPVSNACSGPCQNGTCLQHTQVSEFCLCKPPFMGPTCAKK 249
QY 408 IGYSGPNCDLQDNCSPNCPINGSC-QPS--G-KCICPAGFSGTRCETN 453

RESULT 7
ID PCT-US91-09055-2 STANDARD; PRT; 1480 AA.
XX
XX
AC
XX
XX
DT 01-JAN-1900
XX
DE
DE
CC Sequence 2, Application PC/TUS9109055.
CC Sequence 2, Application PC/TUS9109055
CC GENERAL INFORMATION:
CC APPLICANT: Rothberg, Jonathan Marc and Artavanis-Tsakonas, Spyridon
CC TITLE OF INVENTION: Purified SLIT protein and Sequence Elements Thereof
CC NUMBER OF SEQUENCES: 9

QY 120 FSWPCTFSLVIAWHDNTNNSGNARKLIIQRLIIYQVLEWSEWTKNSESYTSLDYD 179
Db 233 VRVQCAVYNTCTTCFPRDDQFGHYACGSEGGQKCLNGWQGVNCEAICKAGCDPVH 292
QY 180 FRVTCDLNYGSGCAKFCRPRDDSFHSGTSETGEIICLTGWQDYCHIPKCAKGEHGH 239

RESULT 9
ID PCT-US95-11684-2 STANDARD; PRT; 2199 AA.
XX
AC xxxxxx
XX
DT 01-JAN-1900
XX
DE Sequence 2, Application PC/TUS9511684.
XX
XX Sequence 2, Application PC/TUS9511684
XX
CC GENERAL INFORMATION:
CC APPLICANT: THE SCRIPPS RESEARCH INSTITUTE
CC TITLE OF INVENTION: CYTOTACTIN DERIVATIVES THAT STIMULATE
CC TITLE OF INVENTION: ATTACHMENT AND NEURITE OUTGROWTH, AND METHODS OF MAKING
CC TITLE OF INVENTION: AND USING SAME
CC NUMBER OF SEQUENCES: 28
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: The Scripps Research Institute, Office of
CC ADDRESSEE: Patent Counsel
CC STREET: 10666 North Torrey Pines Road, TPC 8
CC CITY: La Jolla
CC STATE: CA
CC COUNTRY: USA
CC ZIP: 92037
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US95/11684
CC FILING DATE: 14-SEP-1995
CC CLASSIFICATION:
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/308,359
CC FILING DATE: 16-SEP-1994
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Logan, April C.
CC REGISTRATION NUMBER: 33,950
CC REFERENCE/DOCKET NUMBER: BEC0019P
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 619-554-2937
CC TELEFAX: 619-554-6312
CC INFORMATION FOR SEQ ID NO: 2:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 2199 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 2199 AA; 240716 MW; 25336188 CN;

Query Match 8.0%; Score 493; DB 2; Length 2199;
Best Local Similarity 33.2%; Pred. No. 3.86e-29;
Matches 130; Conservative 62; Mismatches 154; Indels 46; Gaps 37;

Db 207 CDDGTGDCSOLACP-SDCNDQKGVN-GVICIFBGAAD-CSREICPVPCSEEHGTVCV 263
QY 184 CDLYNYGSGCAKFCRPRDDSFHSGTSETGEIICLTGWQDYCHIPKCAKGC--EHGCD 241
Db 264 D-GLCVCHDGFAGDDCNKPLCLN-NCYNRGRCVEN-ECVCDGEGTGDGDCSELICPNDCFD 320
QY 242 KPNQCVCOLGWKALGALNCECV-LEPNICIIH-GTCNKPWTCICNEGGLYCNDL--NYCTN 297
Db 321 RGRCLNG-TCY-CEEGFTGDCGKPTCPHACTQG-RCEEGQCVCDGFGAGVDCSEKRCP 377

QY 298 HRPCKNGGTCFNTGEGLYTCKCAPGYSGDDCENEIYSCDADVNPCNGQGTCTIDEPHTKTG 357
Db 378 ADCH-NRGRCVDRCECDGFTGAD--CGELKCPNCGSHGRGVNG-QCVCDGEGYTGEDC 433
QY 358 KYCHCANGWSGK-MCE-EKVLTCSDRKPCHQICRNVRPGLGSRGQGYQCECPYGSPPNC 415
Db 434 S-OL-RC-PNDCHSRGR-VGKCVCEQGFKYDC-SDMS-C-PNDCHQHGRC---VNG- 482
QY 416 DIQLDMCSPNCPINGSCQPSGKICIPAGFSGRCTETNIDDCLGHCQENGCTCIDMVNQY 475
Db 483 MCVCDDGYTGEDCRDR-Q-CP-RDCSNRGLCYD---G-QCVCDGEGTGPDC-ELS-CPN 533
QY 476 RCQCVPFHGHGTHCSSKVDLCLIRPCANGGTCLNLNDYQCTCRAGFTGKDCSVDIDECSS 535
Db 534 D-CHGOGRC---VNG-QCVCHGEGFMKDKCEQ 560
QY 536 GPCHNGGTGMNRYNSPECVANGFRGKQCDDE 567

RESULT 10
ID PCT-US95-11684-4 STANDARD; PRT; 1810 AA.
XX
AC xxxxxx
XX
DT 01-JAN-1900
XX
DE Sequence 4, Application PC/TUS9511684.
XX
XX Sequence 4, Application PC/TUS9511684
XX
CC GENERAL INFORMATION:
CC APPLICANT: THE SCRIPPS RESEARCH INSTITUTE
CC TITLE OF INVENTION: CYTOTACTIN DERIVATIVES THAT STIMULATE
CC TITLE OF INVENTION: ATTACHMENT AND NEURITE OUTGROWTH, AND METHODS OF MAKING
CC NUMBER OF SEQUENCES: 28
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: The Scripps Research Institute, Office of
CC ADDRESSEE: Patent Counsel
CC STREET: 10666 North Torrey Pines Road, TPC 8
CC CITY: La Jolla
CC STATE: CA
CC COUNTRY: USA
CC ZIP: 92037
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US95/11684
CC FILING DATE: 14-SEP-1995
CC CLASSIFICATION:
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/308,359
CC FILING DATE: 16-SEP-1994
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Logan, April C.
CC REGISTRATION NUMBER: 33,950
CC REFERENCE/DOCKET NUMBER: BEC0019P
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 619-554-2937
CC TELEFAX: 619-554-6312
CC INFORMATION FOR SEQ ID NO: 4:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1810 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 1810 AA; 199304 MW; 17319263 CN;

Query Match 7.6%; Score 470; DB 2; Length 1810;
Best Local Similarity 34.2%; Pred. No. 2.69e-27;
Matches 135; Conservative 61; Mismatches 145; Indels 54; Gaps 40;

Db	178	CEIMKKGNCSEPRAC-PNN-CLNGLCLVRA-KICCEGFFIGEOSQARCSIDCNDGKCV	234
Oy	184	CDLNTYGGGCKK-FCRPRDDSFSGSTSEGEIILCLTGMQGDYCHIDPKAKGC-EHGHD	241
Db	235	D-GVOCVEEGYATGPDGCE-ELCPHGCGIHGRVCGR-CVCEHEGFTGEDCNEPL-CPNN-	288
Oy	242	KPNQCVQLGKKKALCNECVLEPN-C-IHGCKNPMWICIEBGGGLXCMODLNTYCTNHR	299
Db	289	-CHNHRGVDN-E---CVCDGEGYTEDGCEGLICPNDCFRGRICNG-TCFCEEGYTGED	341
Oy	300	PCKNNGTCFNTGEGLYTCKCAPXSGDDCENEIYSCDA-DVNPQNGGTCT-DEPHKTG	357
Db	342	CGELTCEPNNCNGNRCENGELCVCHEGFPVGDGDSQKRCRPTCNNGRCVND-KVCYCHGL	400
Oy	358	Y-KCHCAAGWSGK-MCEKVTLYTCD---KPCMHGILRNMRPGLSGKSGDYQCECPRTGYS	411
Db	401	GEDCG-EL-RC-PNDCHHRGRCI-NGQVCDEGFIAGEDG-ELR-C-PNCCQDGRGRCIN-	452
Oy	412	GPNCDDLDLDNCSPPNCLINGGSCQSGKACICPAAGISGTRCENIINDDCLGHCCENGCTICDM	471
Db	453	-GQ-CECEHEGFIGEDGE-LR-CP-NDGNSHGRVFN--G-QCVCDEGYTGEDG-EL-	500
Oy	472	VNOYRCQCVLPFGFHGHCSSKVDLILRCPAAGGTCLMLNDYQCTCRAGFTGKDCSVYID	531
Db	501	RCPNND-CHNHRG---YEG-KVCVDNFGMEGDGE	530
Oy	532	ECSSGPCHNNGTGMNRVNSFEVCANBERGKQDCE	566

RESULT 11
 ID US-08-264-534-4 STANDARD; PRT; 199 AA.
 AC xxxxxx
 XX 01-JAN-1900
 DT
 XX
 DE Sequence 4, Application US/08264534.
 XX
 CC Sequence 4, Application US/08264534
 CC Patent No. 5648464
 CC GENERAL INFORMATION:
 CC APPLICANT: Artaevanis-Tsakonas, Spyridon et al.
 CC TITLE OF INVENTION: Human No. 5648464ch And delta, Binding Domains
 CC TITLE OF INVENTION: In Topoprychmic Proteins, And Methods Based Thereon
 CC
 CC NUMBER OF SEQUENCES: 34
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: Pennie & Edmonds
 CC STREET: 1155 Avenue of the Americas
 CC CITY: New York
 CC STATE: New York
 CC COUNTRY: U.S.A.
 CC ZIP: 10036
 CC
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Floppy disk
 CC COMPUTER: IBM PC compatible
 CC OPERATING SYSTEM: PC-DOS/MS-DOS
 CC SOFTWARE: Patentin Release #1.0, Version #1.25
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/08/264,534
 CC FILING DATE:
 CC CLASSIFICATION: 435
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: US 07/695,189
 CC FILING DATE: 03-MAY-1991
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Mistrock, S. Leslie
 CC REGISTRATION NUMBER: 18,872
 CC REFERENCE/DOCKET NUMBER: 7326-004
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: 212 790-9090
 CC TELEFAX: 212 8698864/9741
 CC TELEX: 66141 PENNIE
 CC

Cc	CC	SEQUENCE CHARACTERISTICS:
Cc	CC	LENGTH: 199 amino acids
Cc	CC	TYPE: amino acid
Cc	CC	STRANDEDNESS: single
Cc	CC	TOPOLOGY: unknown
Cc	CC	MOLECULE TYPE: peptide
SQ	SEQUENCE	199 AA; 21952 MW; 208406 CN;
Query Match	7.3%;	Score 452; DB 1; Length 199;
Best Local Similarity	37.0%;	Pred. No. 7.36e-26;
Matches	77; Conservative	51; Mismatches 66; Indels 14; Gaps 9.
Dd	1	GNELEILEISNTSHLNGYCCGMPAELRATKIGSPCTTARLCLKREVOTTEGASI 60
Oy	24	GSFEKLRYTFENDNRGRDRCSCGESDGATGKLG-S-CITRRVCLKHQAID-TT- 79
Dd	61	STGCSFGNATTKILIGSSFEVLSDP-----G-VGAIVLPETFRWTKSFLLQA-IDMYN 112
Oy	80	SQ-CYEGDVLTPIIGENSVNLTDQRPNKGFTNPIDPFPSFSGFFSLVEAMHDTNN 138
Dd	113	TSTPAEKLIEETISGYILP-SEPKTTLDHIGNNAITYRVQCAVTYYNTTCITFCR 171
Oy	139	SGNATNKILLIQLRLVVOGVLEVSSEMKTNSKSQYTSLIEDYDFRTCDLNYGSGCAKFCR 198
Dd	172	PRDDQFGHYAGCSBQKRLCIMGNVNC 199
Oy	199	PRDDSFGHSTCSFETGEIICLTGMWGDDYC 226

RESULT	12			
ID	PCT-US95-02251-3	STANDARD;	PRT;	1251 AA.
XX	xxxxxx			
XX	01-JAN-1900			
XX	Sequence 3, Application PC/TUS9502251.			
DE				
CC	Sequence 3, Application PC/TUS9502251			
CC	GENERAL INFORMATION:			
CC	APPLICANT:			
CC	TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR STIMULATING BONE			
CC	TITLE OF INVENTION: CELLS			
CC	NUMBER OF SEQUENCES: 18			
CC	CORRESPONDENCE ADDRESSES:			
CC	ADDRESSEE: Arnold, White & Durkee			
CC	STREET: P. O. Box 4433			
CC	CITY: Houston			
CC	STATE: Texas			
CC	COUNTRY: United States of America			
CC	ZIP: 77210			
CC	COMPUTER READABLE FORM:			
CC	MEDIUM TYPE: Floppy disk			
CC	COMPUTER: IBM PC compatible			
CC	OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII			
CC	SOFTWARE: PatentIn Release #1.0, Version			
CC	SOFTWARE: #1.30			
CC	CURRENT APPLICATION DATA:			
CC	APPLICATION NUMBER: PCT/US95/02251			
CC	FILING DATE: CONCURRENTLY HERewith			
CC	CLASSIFICATION:			
CC	PRIOR APPLICATION DATA:			
CC	APPLICATION NUMBER: US 08/316,650			
CC	FILING DATE: 30-SEP-1994			
CC	CLASSIFICATION:			
CC	APPLICATION NUMBER: US 08/199,780			
CC	FILING DATE: 18-FEB-1994			
CC	CLASSIFICATION:			
CC	ATTORNEY/AGENT INFORMATION:			
CC	NAME: Parker, David L.			
CC	REGISTRATION NUMBER: 32,165			
CC	REFERENCE/DOCKET NUMBER: DWIC009P--			

CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (512) 418-3000
CC TELEFAX: (713) 789-2679
CC TELEX: 79-0924
CC INFORMATION FOR SEQ ID NO: 3:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1251 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
SQ SEQUENCE 1251 AA; 134551 MW; 7366256 CN;

Query Match 6.3%; Score 388; DB 2; Length 1251;
Best Local Similarity 32.8%; Pred. No. 8.84e-21;
Matches 117; Conservative 58; Mismatches 129; Indels 53; Gaps 44;

Db 555 DECLNQCIGHCQVPGSDYSDYCHNAGYRSHPHQRYCY-DVNECEAE-PCGPGKGICM 612
QY 258 NECVLEPN-CIHCTCNK-P--WTCICNEG-WG-CL--YCNODLNYCTNHRPCKNG-GTCF 308
Db 613 NTG-GSYNCHNRYRLHVGAGR-SC-VDLNECAKPHLCGDGFGCINFPGHYKNCYPG 669
QY 309 NTGEGLYTCKCAPGSDCCENEIYSCDADVNPQNGGTCIDEPH-TK-TG-YKCHCANG 365
Db 670 YRLK--ASRPICEDIDECRDPSTCPDGKCE--NKPGSFKCIACQPCYRSGGGACR-DV 724
QY 366 WSGKMEKEKVLTCSD-KPCHQ-GICRNVPRGLSGKGQYQC-ECPIGY-S--GPNCDLQL 419
Db 725 NECSEGTTC-SPGWCEKLPQSYRCTCAQIRTRTGRISC-IDVDDCAGKVCODG-ICTN 781
QY 420 DNGSPN-PCINGSCQ--PSG-KCICPAGEFG-T-R--CETNIDDC-LGHQENGGTCTD 470
Db 782 TPQSFQCLSGYHLSRDRSCEIDEDCPAACI-GGDCINTNGSYRCLPGLHRLVGG 840
QY 471 MVNQRQCQVPGFH-G-THCSSK-VDLK-LIRPCANGGTCLNNDYQCTCRAG--FTG- 523
Db 841 RKCKKIDECSDPGLCLPH-ACENLOGSYVCVDEGFTLTQ-DOHCCEVE-QPHH 894
QY 524 KDCSDVIDECSSGP--CHNGGTGMNRVNSPECVANGFRQCKDESYDSTFDAQH 578

RESULT 13
ID PCT-US95-02251-18 STANDARD; PRT; 1833 AA.
XX xxxxxx
XX 01-JAN-1900
DE Sequence 18, Application PC/TUS9502251.
XX
CC Sequence 18, Application PC/TUS9502251
CC GENERAL INFORMATION:
CC APPLICANT:
CC TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR STIMULATING BONE
CC TITLE OF INVENTION: CELLS
CC NUMBER OF SEQUENCES: 18
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Arnold, White & Durkee
CC STREET: P.O. Box 4433
CC CITY: Houston
CC STATE: Texas
CC COUNTRY: United States of America
CC ZIP: 77210
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: .PC-DOS/MS-DOS/ASCII
CC SOFTWARE: PatentIn Release #1.0, Version
CC SOFTWARE: #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US95/02251
CC FILING DATE: CONCURRENTLY HEREWITH
CC CLASSIFICATION:







CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/316,650
CC FILING DATE: 30-SEP-1994
CC CLASSIFICATION:
CC APPLICATION NUMBER: US 08/199,780
CC FILING DATE: 18-FEB-1994
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Parker, David L.
CC REGISTRATION NUMBER: 32,165
CC REFERENCE/DOCKET NUMBER: UMIC009P--
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (512) 418-3000
CC TELEFAX: (713) 789-2679
CC TELEX: 79-0924
CC INFORMATION FOR SEQ ID NO: 18:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1833 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
SQ SEQUENCE 1833 AA; 197927 MW; 15940870 CN;

Query Match 6.3%; Score 388; DB 2; Length 1833;
Best Local Similarity 30.8%; Pred. No. 8.84e-21;
Matches 86; Conservative 54; Mismatches 104; Indels 35; Gaps 33;

Db 870 GTCVSLPVG-YRCVSPGYQLHPSQ-DYCTDDNECMRNPCEGRGVN-S-V-GSYSLC 924
QY 305 GTCFTNCEGLYTKCAPGYSDDDCENEIYSCDAD-V-NPCQNGGTCTDEPHTKTYKCHC 362
Db 925 YPGYT-LVTLDGTQECQDIDECQPGVCSGGRCS-NTEG-SYHCECDRGYIMVRKHGHCQ- 980
QY 363 ANGWSGKMEKEKVLTCSD-KPCHQ-GICRNVPRGLSGKGQYQCCECPIGYS-GP--NCDL 417
Db 981 DINECRHPGTCPDGRCVNSPGSYTCLACERGVYQSGSC-VDVNECLTGCICTHG-RCIN 1038
QY 418 QLDNCS-PNFCINGSCQ-PSG-KCI-CPAGFSG-T-RCETNIDDCGLHQ-CENGGTCTD 470
Db 1039 MEGSFRCSPFGYEVTPDKKGRDVEDCASRASCPTGLCLNTEGFTSCACQSGYWNED 1098
QY 471 MVNQRQCQVPGFHGT-HCSS-K-VDLCLIRPCANGGTCLNNDYQCT-CRAG-TGKD 525
Db 1099 GTACEDLDECAFPQVCPGTG-VCTNVTGVSFCKDCDQGYR 1136
QY 526 CSV--DIDEGS-SGPNCHNGGTGMNRVNSFECV-CANGFR 560

RESULT 14
ID 5177197-30 STANDARD; PRT; 1510 AA.
XX xxxxxx
XX 01-JAN-1900
DE Patent No. 5177197.
XX
CC Patent No. 5177197
CC APPLICANT: KANZAKI, TETSUO; OLOFSSON, ANDERS; MOREN, ANITA;
CC WERNSTEDT, CHRISTER; HELLMAN, ULF; MIYAZONO, KOHEI; CLAESSON-WELSH,
CC LENA; HELDIN, CARL-HENRIK
CC TITLE OF INVENTION: ISOLATED NUCLEOTIDE SEQUENCE EXPRESSING
CC HUMAN TRANSFORMING GROWTH FACTOR-BETAL-BINDING PROTEIN
CC NUMBER OF SEQUENCES: 53
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/07/487,343
CC FILING DATE: 27-FEB-1990
CC SEQ ID NO: 30:
CC LENGTH: 1394
SQ SEQUENCE 1510 AA; 165567 MW; 12055844 CN;

Query Match 6.1%; Score 379; DB 3; Length 1394;
Best Local Similarity 31.8%; Pred. No. 4.53e-20;

```
QY 301 CKNGTCFNTGEGLYTCRCAPGYSGDDCENEIYSCDADVNPQNGGTCIDEPHTKTGYK 360
Db 361 HCANGSGKMCBEKVLTCTSDKPCCHQICRNVRPGLSGKGGYQCECPIGYSGPNCDLQLD 420
QY 361 HCANGSGKMCBEKVLTCTSDKPCCHQICRNVRPGLSGKGGYQCECPIGYSGPNCDLQLD 420
Db 421 NCSPNFCINGGSCQPSGKICPCAGFSGTRCETNIDDCGLHQCENGGTCIDMVNQYRCQCV 480
QY 421 NCSPNFCINGGSCQPSGKICPCAGFSGTRCETNIDDCGLHQCENGGTCIDMVNQYRCQCV 480
Db 481 PGFHGTHCSSKVDLCLIRPCANGGTCCLNNDYQCTCRAGFTGKDCSVDIDECSSGPGCHN 540
QY 481 PGFHGTHCSSKVDLCLIRPCANGGTCCLNNDYQCTCRAGFTGKDCSVDIDECSSGPGCHN 540
Db 541 GGTCTMNRVNSFCVFCANGFRGKQCEDEESYDSVTFDAHQYGATTOARADGLTNAQVVLIAV 600
QY 541 GGTCTMNRVNSFCVFCANGFRGKQCEDEESYDSVTFDAHQYGATTOARADGLTNAQVVLIAV 600
Db 601 FSVAMPPLVAVIAACVVFCKRKRRAQEKDDAEARKONEQNAVATMHHNGSGVGVALASA 660
QY 601 FSVAMPPLVAVIAACVVFCKRKRRAQEKDDAEARKONEQNAVATMHHNGSGVGVALASA 660
Db 661 SLGGKTSNSGLTFDGGNPNIIKNTWDKSVNNICASAAAAAADADECLMTGGYVASV 720
QY 661 SLGGKTSNSGLTFDGGNPNIIKNTWDKSVNNICASAAAAAADADECLMTGGYVASV 720
Db 721 ADNNNANSDFCVAPLQRAKSQKQNTDPTLMHRGSPAGSSAKGASGGGPGAAEGKRISVL 780
QY 721 ADNNNANSDFCVAPLQRAKSQKQNTDPTLMHRGSPAGSSAKGASGGGPGAAEGKRISVL 780
Db 781 GEGSYCSQRWPSLAAGVAGACSSQLMAAASAGAGTAQOORSVVCCTPHM 833
QY 781 GEGSYCSQRWPSLAAGVAGACSSQLMAAASAGAGTAQOORSVVCCTPHM 833

RESULT 2 A31246 #type complete
ENTRY neurogenic protein Delta precursor - fruit fly (Drosophila)
TITLE melanogaster)
ORGANISM #formal_name Drosophila melanogaster
DATE 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change
16-Feb-1997
ACCESSIONS A31246
REFERENCE Kopczynski, C.C.; Alton, A.K.; Fechtel, K.; Kooh, P.J.;
Muskhavitch, M.A.T.
#journal Genes Dev. (1988) 2:1723-1735
#title Delta, a Drosophila neurogenic gene, is transcriptionally
complex and encodes a protein related to blood coagulation
factors and epidermal growth factor of vertebrates.
#cross-references MUID:89196890
#accession A31246
#molecule_type mRNA
#residues 1-832 #label KOP
#cross-references GB:Y00222
GENETICS
#gene FlyBase:D1
#cross-references FlyBase:FBgn0000463
SUMMARY #length 832 #molecular-weight 88943 #checksum 636
Query Match 99.1%; Score 6107; DB 2; Length 832;
Best Local Similarity 98.7%; Pred. No. 0.00e+00;
Matches 822; Conservative 8; Mismatches 2; Indels 1; Gaps 1;
Db 1 MHWIKLLTAFICFTVIVQVHSSGSFELRLKYFSNDHGRDNEGRCCSGESDGTATKCLGS 60
QY 1 MHWIKLLTAFICFTVIVQVHSSGSFELRLKYFSNDHGRDNEGRCCSGESDGTATKCLGS 60
Db 61 CKTRFRLCLKHQATIDTTSQCTYGVITPILGENSVNLTDQAFQNKFTNPQFPFSF 120
QY 61 CKTRFRLCLKHQATIDTTSQCTYGVITPILGENSVNLTDQAFQNKFTNPQFPFSF 120
Db 121 SWPGTFSLVEAHDNNSGNARTNKLTLQRLLVQVLEVSSEWKNKSESQTSLEYDF 180
```

```
QY 121 SWPGTFSLVEAHDNNSGNARTNKLTLQRLLVQVLEVSSEWKNKSESQTSLEYDF 180
Db 181 RVTCDLNYIYSGCAKFCRPRDDSFHSTCSETGEIICLTGWOGDYCHIPKCAKCEHGH 240
QY 181 RVTCDLNYIYSGCAKFCRPRDDSFHSTCSETGEIICLTGWOGDYCHIPKCAKCEHGH 240
Db 241 DRPNQVCQGLGKAGLNCVLEPNCIHGTCNKPTWCICNEGGWGLYCNDLNYCTNHRP 300
QY 241 DRPNQVCQGLGKAGLNCVLEPNCIHGTCNKPTWCICNEGGWGLYCNDLNYCTNHRP 300
Db 301 CKNGTCFNTGEGLYTCRCAPGYSGDDCENEIYSCDADVNPQNGGTCIDEPHTKTGYK 360
QY 301 CKNGTCFNTGEGLYTCRCAPGYSGDDCENEIYSCDADVNPQNGGTCIDEPHTKTGYK 360
Db 361 HCRNGSGKMCBEKVLTCTSDKPCCHQICRNVRPGLSGKGGYQCECPIGYSGPNCDLQLD 420
QY 361 HCRNGSGKMCBEKVLTCTSDKPCCHQICRNVRPGLSGKGGYQCECPIGYSGPNCDLQLD 420
Db 421 NCSPNFCINGGSCQPSGKICPCAGFSGTRCETNIDDCGLHQCENGGTCIDMVNQYRCQCV 480
QY 421 NCSPNFCINGGSCQPSGKICPCAGFSGTRCETNIDDCGLHQCENGGTCIDMVNQYRCQCV 480
Db 481 PGFHGTHCSSKVDLCLIRPCANGGTCCLNNDYQCTCRAGFTGKDCSVDIDECSSGPGCHN 540
QY 481 PGFHGTHCSSKVDLCLIRPCANGGTCCLNNDYQCTCRAGFTGKDCSVDIDECSSGPGCHN 540
Db 541 GGTCTMNRVNSFCVFCANGFRGKQCEDEESYDSVTFDAHQYGATTOARADGLTNAQVVLIAV 600
QY 541 GGTCTMNRVNSFCVFCANGFRGKQCEDEESYDSVTFDAHQYGATTOARADGLTNAQVVLIAV 600
Db 601 FSVAMPPLVAVIAACVVFCKRKRRAQEKDDAEARKONEQNAVATMHHNGSGVGVALASA 660
QY 601 FSVAMPPLVAVIAACVVFCKRKRRAQEKDDAEARKONEQNAVATMHHNGSGVGVALASA 660
Db 661 SMGGKTSNSGLTFDGGNPNIIKNTWDKSVNNICASAAAAAADADECLMTGGYVASV 720
QY 661 SMGGKTSNSGLTFDGGNPNIIKNTWDKSVNNICASAAAAAADADECLMTGGYVASV 720
Db 721 ADNNNANSDFCVAPLQRAKSQKQNTDPTLMHRGSPAGTSAGKASGGGPGAAEGKRISVL 780
QY 721 ADNNNANSDFCVAPLQRAKSQKQNTDPTLMHRGSPAGTSAGKASGGGPGAAEGKRISVL 780
Db 781 GEGSYCSQRWPSLAAGVAGACSSQLMAAASAGAGTD-GTAQOORSVVCCTPHM 832
QY 781 GEGSYCSQRWPSLAAGVAGACSSQLMAAASAGAGTAQOORSVVCCTPHM 833

RESULT 3 S00670 #type complete
ENTRY gene Delta protein precursor - fruit fly (Drosophila)
TITLE melanogaster)
ALTERNATE_NAMES gene D1 protein
ORGANISM #formal_name Drosophila melanogaster
DATE 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change
16-Feb-1997
ACCESSIONS S00670
REFERENCE S00670
#authors Vaessin, H.; Bremer, K.A.; Knust, E.; Campos-Ortega, J.A.
#journal EMBO J. (1987) 6:3431-3440
#title The neurogenic gene Delta of Drosophila melanogaster is
expressed in neurogenic territories and encodes a putative
transmembrane protein with EGF-like repeats.
#accession S00670
#molecule_type mRNA
#residues 1-880 #label VAE
#cross-references EMBL:X06289
GENETICS
#gene Delta
#cross-references FlyBase:FBgn0000463
KEYWORDS transmembrane protein
FEATURE
1-18. #domain signal sequence #status predicted #label SIG\
```


Qy	542	GTCHNRVNSFECVCANGRGQCDEESTDSTFPAHQIGATTQARADGLTNAQVLLAVE	601
Db	181	SVAMPLVAVIACVVECKMR	200
Qy	602	SVAMPLVAVIACVVECKMR	621

RESULT	8
ENTRY	A36666
TITLE	serrate protein precursor - fruit fly (Drosophila melanogaster)
ORGANISM	#format_name Drosophila melanogaster
DATE	19-Apr-1991 #sequence_revision 19-Apr-1991 #text_change 16-Feb-1997

ACCESSIONS A36666
REFERENCE A36666
#authors Fleming, R.J.; Scottgale, T.N.; Diederich, R.J.; Artavanis-Tsakonas, S.
#journal Genes Dev. (1990) 4:2188-2201
#title The gene *Serrate* encodes a putative EGF-like transmembrane protein essential for proper ectodermal development in *Drosophila melanogaster*.
#cross-references MIMD:91099666

```
#accession      A36666
#status         preliminary
#molecule_type mRNA
#residues       1-1404 #label FLE
#cross-references GB:M35759
```

GENETICS
#gene
#cross-references FlyBase:Fggn0004197
FlyBase:Ser
transmembrane protein
#length 1404 #molecular-weight 150245 #checksum 433
SUMMARY

Query Match	20.58;	Score 1266;	DB 2;	Length 1404;
Best Local Similarity	34.88;	Pred. No. 9.53e-231;		
Matches 207; Conservative	143;	Mismatches 199;	Indels 46;	Gaps 35

Db 63 CNLHLI-LILLVHKISAAGNFEELEISNTNSHLUNGCCGMPRALRATYIGCSCT 121
| | | | : : : : | | : : : | | : : : | | : : : | | : : : | | : : : |
Qy 6 CLTAFICFTIVIQ-VHSSGSEELRLKFPNSNDHGDNEGRCSSGESDGTGKCLG-S-CK 62

DB 122 IAFKRLNEIQLLEGGASISVSFGNATINLGGSSFLVSDP-----G-VGALVLPFI 17
 |||:|||||: :: :||: :| :|: | :|: ||:
 QY 63 TRFRVCLKHYQATID-IT-SQ-CYVGDVITPIIGENSVNLTDQAQRFQNKGTNP IQPFPS 11

```

120 FSWPGTFFSLVEAMHDITNNSGNAPRTKLLIQRLVQOVLEVSSEWKTKNSEQITSLEYD 17

```

QY 180 FRVTCDLNYYGSGCAKFCRPRDSFGHSTCSETGEIICLTGMQGDYCHIPKCAKCE--H 23

Db 353 THEPCKHGGTCENTAPDKRYCTCAEGLSGECCETVEHPBC-A-TRPGKRNGGTCTTIXTSNRT 41
Qy 238 GHCDKPNQCQGLGWKGALDCEVLEPNCIHTGTCN-KPWTCICNKGWGLYCNDLNYCT 29

QY 297 NHRPCKNGTGCTENTGEGLYTCKAPGYSDDCCENEIYSCDAVNPCCQNGTCTC IDEPH -T 35

Qy 355 KTG-YK-CH-CAN-GWSGKCEE-KVLTG--SDKPCHQGICNVRPGLG---KQ-G-- 40

Db 471 FTCDCAAGWTGPTCEINIDCAGGPCHEHGTCIDLIGGFRCPCPEWHGADVQYVNECE 53

QY 402 YQCEPIGYSPCNDLQDNCSPNCINGSC -Q-P-SG-KCICPAGFSGTRETINIDDL 458

```

0y 459 GGGCEN-GGTCT--DMVQYR-CQ-CYPGGFSTHSSKYDCL-IRRCANGTGLNAND 512
Db 591 FACICKEGSGVTCANLDDCV-GQCRNGATCIDLVNDYRCACASGFTGRDCEFD 644
   :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
0y 513 YQCTCRAFETGKDCSDVIDECSSPCFHNGTGCMNRVNSFECVCANGFRGKQDEE 567

```

ENTRY	9
RESULT	
TITLE	Sl6148 #type complete gene serrate protein precursor - fruit fly (Drosophila melanogaster) #formal_name Drosophila melanogaster 31-Dec-1991 #sequence_revision 02-Aug-1994 #text_change
ORGANISM	
DATE	

ACCESSIONS	REFERENCE
16-red-199/ S16148; S16878	Thomas, U.; Spelcher, S.A.; Knust, E. Development (1991) 111:749-761
S16148	The Drosophila gene <i>sevrata</i> encodes an EGF-like transmembrane

```
#cross-references MUID:91347903
#accession SL6148
#molecule_type mRNA
##residues 1-1408 ##label THO1
##cross-references EMBL:X56811
REFERENCE
SL6878
```

#authors Thomas, U.
 #submission submitted to the EMBL Data Library, November 1990
 #accession S16878
 ##molecule-type mRNA
 ##residues 1-1351, 'T', 1353-1408 ##label TH02
 #Cross-references EMBL:X56811
 GENETICS

CLASSIFICATION	KEYWORDS	FEATURE
#cross-references FlyBase:FBgn0004197	#superfamily EGF homology	glycoprotein; transmembrane protein

```

1-84      #domain signal sequence #status predicted #label SIG\
85-1408   #product gene serrate protein #status predicted #label
          MAT\
85-1221   #domain extracellular #status predicted #label EX\
          #status non transmembrane #status predicted #label
          #status non transmembrane #status predicted #label

```

495-526	#domain	EGF	homology	#label	EG05\
355-488	#domain	EGF	homology	#status	atypical
355-388	#domain	EGF	homology	#label	EG03\
319-348	#domain	EGF	homology	#label	EG02\
268-310	#domain	EGF	homology	#label	EG01\

	domain	EGF	homology	#status	atypical	#label	EC06\
533-608	domain	EGF	homology	#label	EC07\		
615-645	domain	EGF	homology	#label	EC07\		
652-683	domain	EGF	homology	#label	EC08\		
690-720	domain	EGF	homology	#label	EC09\		

727-796	#domain EGF homology #status atypical #label EG10\
803-834	#domain EGF homology #label EG11\
841-876	#domain EGF homology #label EG12\
883-914	#domain EGF homology #label EG13\

	#domain	Bcf	homology	#label	ES14
921-952	1	0.0	0.0	0	0
997-1060	1	0.0	0.0	0	0
1222-1246	1	0.0	0.0	0	0
1247-1408	1	0.0	0.0	0	0
153-106	1	0.0	0.0	0	0
247-331	1	0.0	0.0	0	0

	#binding_site	carbohydrate (Asn)	#status predicted
1150			
965, 977, 1004, 1030,			
412, 452, 558, 739,			
102, 150, 241, 531,			

SUMMARY	#length 1408	#molecular-weight 150645	#checksum 5466
Query Match	20.5%	Score 1266;	DB 2; Length 1408;
Best Local Similarity	34.8%;	Pred. No. 9,53e-231;	
Matches 207;	Conservative	133;	Mismatches 199;
		Indels 46;	Gaps 35;

Db 67 CNLTALI-LIIIVHKISAGNFEELEIEISNTNSHLNGYCCGMPALRAITIGCSPCT 125


```
QY 6 CLLTAFICFTVIVQ-VHSSGSPELRKLYFSDNRDGRNEGRCCSGSDGATGKCLG-S-CK 62
Db 126 TAFRLCLKEYOTEOGASISTCGSFNATTKILGSSSFVLSDP-----G-VGAIVLPFT 178
QY 63 TRFRVCLRHXYQATID-TT-SQ-CTYGDVITPILGNSVNLTDQARFQNGFTNPIQFPES 119
Db 179 FRWTKSFLLILOA-LDMYNTSPDAERLIEETSYSGVILP-SPWKTLHIGENAITYR 236
QY 120 FSWPCTFSLIVEAWHDNTNNSGARTNKKLIQRLLOVQVLEWSEWTKNKSQYTSLEYD 179
Db 237 VRVOCAVYVNTCTTFCRPRDDQFHYACGSEGOIKLCLNGWQVNCERAEICRAGCDPVH 296
QY 180 FRVTCDLNYGSGCAKFCRPRDSDSGHSTCSETGEIICLTGQGDYCHIPKCAKGE--H 237
Db 297 GKCDRPGCECRPRGRPLCNMYPVPGCKHSGMSAWKVCVDTNWGILCDQDILNFCG 356
QY 238 GHCDRPNQCVCQLGWKALGALNECVLEPNICHTCN-KPWTCICNEGWGLYCNQDLNYCT 296
Db 357 THEPCKHGTCENTAPDKYKRCCTCABGLSGEOEIVEHPC-A-TRPCRNGGTCTLKTNSRT 414
QY 297 NHRPCKNGGTCTENTEGLYTCKAPGYSGGDCENEIYSCDADVNPONGGTC-IDEPH-T 354
Db 415 QAQVYRTSHGRNMGPRVRRSSMSRLDHLRPEGOALNGSSSGLVSLAGSLQLOQLAPD 474
QY 355 KTG-YK-CH-CAN-GWSGKMCCE-KVLTG--SDKPCHQGICRNVRLGSG---KGQ-G-- 401
Db 475 FTCDCAAGWTGPTCEINIDECAGGCEHGGTCDILIGFRCECPPEWHGVDVQVDVNECE 534
QY 402 YQCEPIGYSGPNCIDLQDNCSPNCGINGSC-QP-SG-KCICPAGFSGTRCETNIDDCIL 458
Db 535 APHSAGIAANALLTTATAIIGSNLSSTALLAALTSVASTSLAIGPCINAKECRNPQGS 594
QY 459 GHQCN-GGTCT--DMVNOYR-CQ-CVPGFHTGTHCSKVDLCL-IRPCANGGTCLNND 512
Db 595 FACICKEGWGVTCAENLDVCY-GQCRNGATCIDLVNDYRCACASGFTGRDCETD 648
QY 513 YOCTCRAGFTGDCSDVIDECSSGPGCHNGGTGMNRVNSFEVCANGFRGKQCDDEE 567

RESULT 10
ENTRY #type complete
TITLE transmembrane protein precursor - zebra fish
ORGANISM #formal_name Brachydanio rerio #common_name zebra fish
DATE 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 12-Sep-1997
ACCESSIONS S42612
REFERENCE S42612
#authors Bierkamp, C.; Campos-Ortega, J.A.
#journal Mech. Dev. (1993) 43:87-100
#title A zebrafish homologue of the Drosophila neurogenic gene Notch and its pattern of transcription during early embryogenesis.
#accession S42612
#status preliminary
#molecule_type mRNA
#residues 1-2437 #label BIE
#cross-references EMBL:X69088; NID:g433866; PID:g433867
CLASSIFICATION #superfamily unassigned ankyrin repeat proteins; ankyrin repeat homology
FEATURE
1915-1947 #domain ankyrin repeat homology #label AN1\
1948-1980 #domain ankyrin repeat homology #label AN2\
1982-2014 #domain ankyrin repeat homology #label AN3\
2048-2080 #domain ankyrin repeat homology #label AN5
SUMMARY #length 2437 #molecular-weight 262306 #checksum 4021
Query Match 16.0%; Score 987; DB 2; Length 2437;
Best Local Similarity 39.1%; Pred. No. 4,40e-172;
Matches 154; Conservative 82; Mismatches 124; Indels 34; Gaps 23;
Db 643 CKRKPDCYKCIDKINGECVCEPGYSGSMCNINIDCALNPNCHNGGTCLDGVSFTCLC 702
QY 197 CRRPDDSGHSTCSETG-EIICLTGQGDYCHIP--KCA-KGCEHG-HC-DKPNQ--CVC 248
```

```
Db 703 PDGFRDATCLSQHNECSSNP-CIHGSCLOQINSYRCVCEAGWGRNCDININECLSN-PC 760
QY 249 OLGWKGALC-----NECVLEPNCHIGTC-N--KPWTICIEGWGGLYCNODLNTCTHRRPC 301
Db 761 VNGGTCKDWTSG-YLCTCRAGFSGPNCMNINPC-AS-NPCLNOGSCIDD--V-AGFKCN 814
QY 302 KNGGTCTNTGEGLYTCKCAPYSGDDCENEIYSCDADVNPQNGGTCIDEPHTTKTYKCH 361
Db 815 CMLPYTGEVCEN-VLAPCSPPRCKNGVCRE-SEDF----QSFSCNCPAGWQGTCEVDI 868
QY 362 CANGWSGKMCERKVL-CSDKPCHQG-ICRNVRLGSKGQGYQCECPIGYSPNCDLQL 419
Db 869 NECVRNPCTNGVCENLRGGFQCRNPGTGCALCENIDDCENPNCSNGVGVQDRVNGFV 928
QY 420 DNGSPNPNNGGSCQ--PSG-KCICPAGFSGTRCETNIDDCLEGHQCENGGTCIDMYNQYR 476
Db 929 CVCCLAGFRCEACEDIDECVSPACRNGGNTDCVNSVYTCSPAGFSGGINCEINTPCTES 988
QY 477 QCQVPGFHTGTHCSKVDLCLIRPCANGGTCLNLDNYOCTCRAGFTGKDCSVIDECSSG 536
Db 989 SCFNGGTCVDGISSFCVCLPFGFTGNYCQHDVNE 1022
QY 537 PCHNGGTGMNRVNSFEVCANGFRGKQCDDESYD 570

RESULT 11
ENTRY #type fragment
TITLE Motch B protein - mouse (fragment)
ORGANISM #formal_name Mus musculus #common_name house mouse
DATE 21-Jan-1994 #sequence_revision 05-Jan-1996 #text_change 31-Oct-1997
ACCESSIONS A49175; S32113
REFERENCE A49175; PH1570;
#authors Lardelli, M.; Lendahl, U.
#journal Exp. Cell Res. (1993) 204:364-372
#title Motch A and Motch B--two mouse Notch homologues coexpressed in a wide variety of tissues.
#cross-references MUID:93178563
#accession A49175
#status preliminary; nucleic acid sequence not shown
#molecule_type mRNA
#residues 1-1203 #label LAR
#cross-references EMBL:X68279; NID:g287989; PID:g287990
#note sequence extracted from NCBI backbone (NCBIP:126158)
COMMENT This protein has many EGF repeats and lin-12/Notch repeats.
COMMENT This protein is one of the neurogenic proteins controlling the decision between ectodermal and neural fate for cells in the early embryo.
CLASSIFICATION #superfamily unassigned ankyrin repeat proteins; ankyrin repeat homology
SUMMARY #length 1203 #checksum 910
Query Match 15.7%; Score 970; DB 2; Length 1203;
Best Local Similarity 40.9%; Pred. No. 1.57e-168;
Matches 146; Conservative 74; Mismatches 108; Indels 29; Gaps 17;
Db 401 YSQVNECLSNPCITHNGTGGLSGYKCLCDAGWGVNCEYDKNECLSNP-CQNGGTCNNLV 459
QY 225 YCHIPKC-ARGCEHGHG-DK-PN-QCVQLGWKALGJ-----NECVLEPNCHIG-TCKN-- 273
Db 460 NGVRCYCKKGFYNGVNCVNIDECASN-PCLNQGTCTFDDVSG-YTCHMLPYTKGNQOTVL 517
QY 274 -PWTCTICNEGWGLYCNQDLNYCTNHRPCANGGTCFNTGEGLYTCKCAPYSGDDCENEI 332
Db 518 APCSP--NPCEAAVCKEAPNFES-FSCLCAPQWQGRKCTVDYDECISKPCMNGVCHNT 574
QY 333 YSCDADVNPQNGGTCIDEPHTTKYKCHCANGWSGKMCERKVLTCSDKPC-HQGICRNV 391
Db 575 Q--GS-----YVCECPFGSGMGCDEEDINDCLANPCONGSGSCVDHVNFTFSCQCHPGFIGD 627
QY 575 Q--GS-----YVCECPFGSGMGCDEEDINDCLANPCONGSGSCVDHVNFTFSCQCHPGFIGD 627
```

```

OY      392  RPLGSKGQGYOCBCPIGYSGNCDLQDNCSPNFCINGGSC--QPSG-KCICPAGFSGT 448
Db      628  KCOTDMNECTSEPCKNNGGTCSDYNSYCTCTCPAGFHGVHCNNIDECESSCFNGGTCVD 687
OY      449  RCEINIDDCJGHQCEGTCIDMNYQRCQCVPGFHGHGSSKVDLCIRPCANGTCLN 508
Db      688  GINSFSCDCEVGFPGFPCLDHINECSSNPNCLNAGTCVGLGTYRCLICPLGYGKNCQ 744
OY      509  LNDNYOCTCAAGFTGKCKSCVDIDECSSGPCHNGGICMNRVNSFEVCANGFRGKQCD 565

RESULT  12
ENTRY   A49128      #type complete
TITLE   cell-fate determining gene Notch2 protein - rat
ORGANISM #formal_name Rattus norvegicus #common_name Norway rat
DATE    21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change
        12-Sep-1997

ACCESSIONS
REFERENCE #authors      Weinmaster, G.; Roberts, V.J.; Lemke, G.
          #journal      Development (1992) 116:931-941
          #title        Notch2: a second mammalian Notch gene.
          #cross-references MUID:93202015
          #accession     A49128
          #status        preliminary: not compared with conceptual translation
          #molecule_type mRNA
          #residues      1-2471 ##label WEI
          #experimental_source Schwann cell
          #note           sequence extracted from NCBI backbone (NCBIP:127811)
          #superfamily_unassigned ankyrin repeat proteins; ankyrin
          #repeat_homology repeat homology

FEATURE
1876-1908 #domain ankyrin repeat homology #label AN1\
1909-1941 #domain ankyrin repeat homology #label AN2\
1943-1975 #domain ankyrin repeat homology #label AN3\
2009-2041 #domain ankyrin repeat homology #label AN5
SUMMARY   #length 2471 #molecular_weight 265367 #checksum 5929

Query Match      15.7%; Score 967; DB 2; Length 2471;
Best Local Similarity 44.1%; Pred. No. 6,656-168;
Matches 139; Conservative 64; Mismatches 90; Indels 22; Gaps 16;

Db      722  NEC-LSPCHGNGTGLSGKCLCDAGVAVINCEVDKNECLSN-PCONGTCNNLVNG- 778
OY      258  NECVLFPNCIHGTCKN--P-WTCICNGWGLYCNODLNTNRRPCNNGTCFNTGGL 314
Db      779  YRCTCKGFGYNCQVAVIDEC-AS-NPCLNNGTCLD--V-SGYTCHMLPYTGKNCQ-T 832
OY      315  YTKCARGYSGDCEENITYSCDADVPFCQNGGTCIDEHRTYTKCHCANGSGKMEEK 374
Db      833  VLAFCSPNCEMAVCKEA-PNF---ESFTCLCAPGMQGRCTVDYDECVSKPCMNNGI 887
OY      375  VLA-CSDKPC-HQICINVRGLSKGQGYCCECPIGSGPNCQLQDNCSPNFCINGGS 432
Db      888  CHHTQGSYMECEPPGEGSCMDCEEDINDCLANPCQNGSCVDKVTFTSLCLPGFYGDKCQ 947
OY      433  CQPS-G--KCICPAGFSGTGTNIDDCJGHQCEGTCIDMNYQRCQCVPGFHGHGTCSS 489
Db      948  TDNNECSECECKNGGTCSDYVNSYCTCPAGFHGVHCNNIDECESSCFNGGTCYDGIN 1007
OY      490  SKVDLCIRPCANGTCLNNDYCTCRAGFTGKDKSDVIDECSSGPCHNGGTCMNRVN 549
Db      1008  SFSCLCPVGTGPGFC 1022
OY      550  SFEVCVANGFRGKQCD 564

RESULT  13
ENTRY   A40136      #type complete
TITLE   fibropellin Ia - sea urchin (Strongylocentrotus purpuratus)
ALTERNATE_NAMES epidermal growth factor homolog precursor
CONTAINS        alternatively spliced fibropellin Ib (EGIT)
ORGANISM        #formal_name Strongylocentrotus purpuratus #common_name

```

```

DATE      13-May-1992 #sequence_revision 17-Sep-1997 #text_change
ACCESSIONS A40136; B40136; C40136; A29316; A43131
REFERENCE
#authors      Delgado-Illó-Reynoso, M.G.; Rollo, D.R.; Hursh, D.A.; Raff,
          #journal      J. Mol. Evol. (1989) 29:314-327
          #title        Structural analysis of the uegf gene in the sea urchin
          #cross-references GB:X17530; NID:g10225; PID:g667061
          #accession     B40136
          #status        preliminary: not compared with conceptual translation
          #molecule_type DNA
          #residues      181-251,329-370,'R','372-408','RA','411-441 ##label DE2
          #accession     C40136
          #status        preliminary: not compared with conceptual translation
          #molecule_type DNA
          #residues      'K','747-821,898-978 ##label DE3
          #accession     A29316
          #authors      Hursh, D.A.; Andrews, M.E.; Raff, R.A.
          #journal      Science (1987) 237:1487-1490
          #title        A sea urchin gene encodes a polypeptide homologous to
          #cross-references MUID:87319677
          #accession     A29316
          #status        preliminary
          #molecule_type mRNA
          #residues      'S','280-481,786-1064 ##label HUR
          #cross-references GB:M17421; NID:g161474; PID:g552260
          #accession     A43131
          #authors      Hunt, L.T.; Barker, W.C.
          #journal      FASEB J. (1989) 3:1760-1764
          #title        Avidin-like domain in an epidermal growth factor homolog from
          #cross-references MUID:89196806
          #contents      annotation
          #note           EGF homology repeats 10-17 are spliced out in the short form
          #superfamily_c1r/c1s repeat homology; EGF homology

CLASSIFICATION
1-19 #domain signal sequence #status predicted #label SIG\
20-1064 #product fibropellin I #status predicted #label FIB\
20-56 #domain EGF homology #label EGF1\
57-175 #domain C1r/C1s repeat homology #label CSR\
176-213 #domain EGF homology #label EGF2\
214-251 #domain EGF homology #label EGF3\
252-289 #domain EGF homology #label EGF4\
290-327 #domain EGF homology #label EGF5\
328-365 #domain EGF homology #label EGF6\
366-403 #domain EGF homology #label EGF7\
404-441 #domain EGF homology #label EGF8\
442-479 #domain EGF homology #label EGF9\
480-517 #domain EGF homology #label EGF10\
518-555 #domain EGF homology #label EGF11\
556-593 #domain EGF homology #label EGF12\
594-631 #domain EGF homology #label EGF13\
632-669 #domain EGF homology #label EGF14\
670-707 #domain EGF homology #label EGF15\
708-745 #domain EGF homology #label EGF16\
746-783 #domain EGF homology #label EGF17\
784-821 #domain EGF homology #label EGF18\
822-859 #domain EGF homology #label EGF19\
860-897 #domain EGF homology #label EGF20\
898-935 #domain EGF homology #label EGF21\
936-1064 #region avidin-like\

```

23-34,28-43,45-54,
62-88,180-191,
185-200,202-211,
218-229,223-238,
240-249,256-267,
261-276,278-287,
294-305,329-314,
316-325,332-343,
337-352,354-363,
370-381,375-390,
392-401,408-419,
413-428,430-439,
446-457,451-466,
468-477,484-495,
489-504,506-515,
522-533,527-542,
544-553,560-571,
565-580,582-591,
598-609,603-618,
620-629,636-647,
641-656,658-667,
674-685,679-694,
696-705,712-723,
717-732,734-743,
750-761,755-770,
772-781,788-799,
793-808,810-819,
826-837,831-846,
848-857,864-875,
869-884,886-895,
902-913,907-922,
924-933

disulfide_bonds #status predicted
#length 1064 #molecular-weight 112072 #checksum 303

Query Match 15.6%; Score 959; DB 2; Length 1064;
Best Local Similarity 40.4%; Pred. No. 3,11e-166;
Matches 129; Conservative 71; Mismatches 99; Indels 20; Gaps 16;
Db 261 CLNGGTCVGV-WN-PECTCLAGTGVRCVNVNIDECASA-PCONGGICIBGING-YTCSC 316
QY 260 CVLEPNICIRGCTKPKWTCICNEGWGLYCNOQLNYCTNHRPKNGGTCFNTGGLYCKC 319
Db 317 PLGFSGDNENDDCESS-I-PCLINGTCVD--LVNA-YMCVAPGWGTPCADNIDEC 371
QY 320 APGSGDDCENEIYSCDADVNPQNGGTCTDEPHTKTYGKCHCANGSGWKCEKVLTC 379
Db 372 SAPCQNGVC--I-DGV--NG--YMCDCQPGYTGTCTETDIDECARPPCQNGDCVDGVN 424
QY 380 DRPCHOG-ICRNVRLGSGGQYQCEPIGSGPNCQLQDNCSPNCPINGGSC-QP-S 436
Db 425 GYVCICAPGFDGLNCNNIDECASRPCQNGAVCVGVNGVFCVTCAGYTGVLCTDINEC 484
QY 437 G-KCICPAGFSGTRCETNIDCLGHQCENGGTCTDMVNOYRCQCPVGFHGTGHCSSKYDL 495
Db 485 ASMPCLNGVCTDLVNGYICTCAAGFEGTNCETDIDECASFPQNGATCTDOVNGYVCTC 544
QY 496 LTRPCANGGTCLNNNDYQCTCRAGTGTGKDCSVVIDECSSGPGCHNGGTCMNRVNSFCVC 555
Db 545 VFGYTGVLCTDINECAF 563
QY 556 ANGFRGKQCDSESYDVTWF 574

RESULT 14
ENTRY A35844 #type complete
TITLE Xotch protein - African clawed frog
ORGANISM #formal_name Xenopus laevis #common_name African clawed frog
DATE 12-Oct-1990 #sequence_revision 12-Oct-1990 #text_change 12-Sep-1997
ACCESSIONS A35844
REFERENCE A35844
#authors Coffman, C.; Harris, W.; Kintner, C.
#journal Science (1990) 249:1438-1441

#title Xotch, the Xenopus homolog of Drosophila notch.
#cross-references MUID:90385285
#accession A35844
#status preliminary; nucleic acid sequence not shown; not compared with conceptual translation
#molecule_type mRNA
#residues 1-2524 #label COF
CLASSIFICATION #superfamily unassigned ankyrin repeat proteins; ankyrin repeat homology
transmembrane protein
KEYWORDS
FEATURE
1924-1956 #domain ankyrin repeat homology #label AN1
1957-1989 #domain ankyrin repeat homology #label AN2
1991-2023 #domain ankyrin repeat homology #label AN3
2057-2089 #domain ankyrin repeat homology #label AN5
SUMMARY #length 2524 #molecular-weight 274931 #checksum 9441

Query Match 15.5%; Score 953; DB 2; Length 2524;
Best Local Similarity 36.0%; Pred. No. 5.56e-165;
Matches 134; Conservative 95; Mismatches 113; Indels 30; Gaps 24;
Db 222 CNPSPCLNGTCTCRQTD-DTSYDCTCLPFGSQNCEENIDDCPSNCRNGTCTVDGVNTYN 280
QY 226 CHIPKCAKG--CEHGCHDKPQCVCQLGWKALCNECVLE-P--NCIHG-TC-N--KPWT 276
Db 281 CQCPDWTGYCTEDVDECOLMPNACONGTCHNTYGG-YNCVYNGWTGDCSENIDDC 339
QY 277 CINCNEGGLYCNOQLNYCT-NHRPKNGGTCFNTGGLYTCAPGSGDDCENEIYSC 335
Db 340 -ANA-ACHGATC-HD-RVASFY-CECPHRTGLLCHLDN-ACISNPNCEGSGNC-DTNP- 391
QY 336 DADVNPQNGGTCTDEPHTKTYGKCHCANGSGWKCEKVLTCSDKPCHOGI-CRNVRP 394
Db 392 VNGKA--I-CTCPPGYTPACNNNDVDECSLGANPCERGGRTNTLGSFQCNCPQGYAGPR 448
QY 395 LSGKGQGYCECPGIGSGPNCQLQDNC--PNPCINGSGQPS-G--KCICPAGFSGTR 449
Db 449 CEIDVNECLNPNQNDSTCLDQIGEFQICMPGYEGLYCETNIDECASNCLHNGKIDK 508
QY 450 CETNIDCLGHQCENGGTCTDMVNOYRCQCPVGFHGTGHCSSKYDLCLIRPANGGTCLNL 509
Db 509 INFRCDCPTGFSGLCOHDFDECTSTPKNGAKCLDGPNSYTCQCTEGTGRHCEODIN 568
QY 510 NNHYQCTCRAGTGTGKDCSVVIDECSSGPGCHNGGTCMNRVNSFCVCANGFRGKQCDSE 569
Db 569 ECIP-DPCHYGT 579
QY 570 DSVTFDAHYGA 581

RESULT 15
ENTRY A24420 #type complete
TITLE 'notch protein - fruit fly (Drosophila melanogaster)
ORGANISM #formal_name Drosophila melanogaster
DATE 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 20-Mar-1998
ACCESSIONS A24420; A24768
REFERENCE A24420
#authors Kidd, S.; Kelley, M.R.; Young, M.W.
#journal Mol. Cell. Biol. (1986) 6:3094-3108
#cross-references MUID:87064624
#accession A24420
#molecule_type DNA
#residues 1-2703 #label KID
#cross-references GB:K03508; NID:9157991; PID:9157993
REFERENCE A24768
#authors Wharton, K.A.; Johansen, K.M.; Xu, T.; Artavanis-Tsakonas, S.
#journal Cell (1985) 43:567-581
#cross-references MUID:86079539
#accession A24768
#molecule_type mRNA
#residues 1-48, 'I', 50-118, 'R', 120-230, 'I', 232-256, 'N', 258-266, 'A', 268-872, 'R', 874-958, 'R', 960-1970, 'FH', 1973-2256, 'G',

 W P E E H (TM)

Release 3.1a John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution Rights by Oxford Molecular Ltd

Msrch_pp protein - protein database search, using Smith-Waterman algorithm
 Run on: Fri Jun 12 11:53:42 1998; MasPar time 21.43 Seconds
 974.872 Million cell updates/sec
 Tabular output not generated.

Title: >US-08-083-590A-2
 Description: (1-833) from US08083590A.pep (1 of 2)
 Perfect Score: 6164
 Sequence: 1 MHYIKCLLRAFCFTYIVQV.....GSGNGRAGQQRSVYCGTPHM 833

Scoring table:
 PAM 150
 Gap 11

Searched: 69111 seqs, 25083644 residues

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database: swiss-prot35
 1:swiss1

Statistics: Mean 50.001; Variance 79.902; scale 0.626

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	6045	98.1	880	DL_DROME	NEUROGENIC LOCUS DELTA	0.00e+00
2	1826	29.6	722	DL1_MOUSE	DELTA-LIKE PROTEIN 1 P	0.00e+00
3	1819	29.5	714	DL1_RAT	DELTA-LIKE PROTEIN 1 P	0.00e+00
4	1266	20.5	1408	SERR_DROME	SERRATE PROTEIN PRECUR	5.51e-285
5	987	16.0	2437	NOTC_BRARE	NEUROGENIC LOCUS NOTCH	1.98e-212
6	959	15.6	1064	FBP1_STRPU	FIBROPELIN 1 PRECURSO	3.37e-205
7	956	15.5	2524	NOTC_XENLA	NEUROGENIC LOCUS NOTCH	2.01e-204
8	951	15.4	2703	NOTC_DROME	NEUROGENIC LOCUS NOTCH	3.32e-197
9	928	15.1	2531	NOTC1_MOUSE	NEUROGENIC LOCUS NOTCH	3.91e-203
10	920	14.9	2531	NOTC1_RAT	NEUROGENIC LOCUS NOTCH	1.43e-192
11	910	14.8	2444	NOTC1_HUMAN	NEUROGENIC LOCUS NOTCH	1.75e-189
12	898	14.6	2318	NOTC3_MOUSE	FIBROPELIN C PRECURSO	7.43e-185
13	881	14.3	570	FBP3_STRPU	NEUROGENIC LOCUS NOTCH	4.10e-125
14	798	12.9	1964	CRB_MOUSE	CRUMBS PROTEIN PRECURS	1.93e-155
15	765	12.4	2139	DLK_HUMAN	DELTA-LIKE PROTEIN PRE	2.48e-127
16	654	10.6	383	DLK_MOUSE	DELTA-LIKE PROTEIN PRE	2.48e-127
17	654	10.6	383	DLK_MOUSE	DELTA-LIKE PROTEIN PRE	2.48e-127
18	647	10.5	1480	SLIT_DROME	SLIT PROTEIN PRECURSO	4.62e-124
19	641	10.4	1429	GLP1_CAEL	GLP-1 PROTEIN PRECURSO	1.33e-99
20	543	8.8	1295	GLP1_CAEL	GLP-1 PROTEIN PRECURSO	1.33e-99
21	493	8.0	2199	APX1_HUMAN	APX-1 PROTEIN PRECURSO	2.66e-86
22	489	7.9	515	APX1_CAEL	APX-1 PROTEIN PRECURSO	2.66e-86
23	464	7.5	1808	TENA_CHICK	TENASCIN PRECURSOR (TN	3.40e-80

RESULT	1	STANDARD	PRT	880 AA.
ID	DL_DROME			
AC	P10041;			
DT	01-MAR-1989 (REL. 10, CREATED)			
DT	01-MAR-1989 (REL. 10, LAST SEQUENCE UPDATE)			
DT	01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)			
DE	NEUROGENIC LOCUS DELTA PROTEIN PRECURSOR.			
GN	DL.			
OS	DROSOPHILA MELANOGASTER (FRUIT FLY).			
OC	EUKARYOTA; METAZOA; ARTHROPODA; INSECTA; DIPTERA.			
RA	VAESSIN H., BREMER K.A., KNUST E., CAMPOS-ORTEGA J.A.;			
RA	EMBO J. 6:3431-3440(1987).			
RN	[2]			
RP	SEQUENCE OF 422-621 FROM N.A.			
RX	MEDLINE: 87218537			
RA	KNUST E., DIETRICH U., TEPASS U., BREMER K.A., WEIGEL D., VAESSIN H.,			
RA	CAMPOS-ORTEGA J.A.;			
RL	EMBO J. 6:761-766(1987).			
RN	[3]			
RP	PATTERN OF TRANSCRIPTION.			
RX	MEDLINE: 91209246			
RA	HAENTLIN M., KRAMATSCHEK B., CAMPOS-ORTEGA J.A.;			
RL	DEVELOPMENT 110:905-914(1990).			
CC	- FUNCTION: ESSENTIAL FOR PROPER DIFFERENTIATION OF ECTODERM. DL			
CC	IS REQUIRED FOR THE CORRECT SEPARATION OF NEURAL AND EPIDERMAL			
CC	CELL LINEAGES.			
CC	- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.			
CC	- SEPARATION OF NEUROBLASTS FROM THE ECTODERM INTO THE INNER PART			
CC	OF EMBRYO IS ONE OF THE FIRST STEPS OF CNS DEVELOPMENT IN INSECTS,			
CC	THIS PROCESS IS UNDER CONTROL OF THE NEUROGENIC GENES.			
CC	- NOTCH AND SERRATE MAY INTERACT AT THE PROTEIN LEVEL, IT IS			
CC	CONCEIVABLE THAT THE SERRATE AND DELTA PROTEINS MAY COMPLETE			
CC	FOR BINDING WITH THE NOTCH PROTEIN.			
CC	- SIMILARITY: CONTAINS 9 EGF-LIKE DOMAINS.			
CC	- SIMILARITY: TO DROSOPHILA SERRATE PROTEIN.			
CC	EMBL: X06289; G7853; "			
CC	EMBL: X05140; G929563; "			
DR	PIR: S00670; S00670.			
DR	PIR: A26637; A26637.			
DR	PIR: A26637; A26637.			
DR	HSSP: P00725; ICBH.			
DR	FLYBASE: FBgn0000463; DL.			
DR	PROSITE: PS00010; ASX_HYDROXYL. 3.			
DR	PROSITE: PS00022; EGF_1; 9.			
DR	PROSITE: PS01186; EGF_2; 9.			
DR	PROSITE: PS01187; EGF_CA; 2.			

ALIGNMENTS

KW	DIFFERENTIATION: NEUROGENESIS; REPEAT: TRANSMEMBRANE;	
KW	EGF-LIKE DOMAIN; GLYCOPROTEIN; SIGNAL.	
FT	SIGNAL	18
FT	CHAIN	19
FT	DOMAIN	19
FT	TRANSMEM	654
FT	DOMAIN	677
FT	DOMAIN	678
FT	DOMAIN	227
FT	DOMAIN	256
FT	DOMAIN	291
FT	DOMAIN	331
FT	DOMAIN	374
FT	DOMAIN	418
FT	DOMAIN	453
FT	DOMAIN	491
FT	DOMAIN	529
FT	DOMAIN	565
FT	DISULFID	231
FT	DISULFID	235
FT	DISULFID	248
FT	DISULFID	260
FT	DISULFID	266
FT	DISULFID	279
FT	DISULFID	295
FT	DISULFID	301
FT	DISULFID	319
FT	DISULFID	335
FT	DISULFID	342
FT	DISULFID	362
FT	DISULFID	378
FT	DISULFID	383
FT	DISULFID	406
FT	DISULFID	422
FT	DISULFID	427
FT	DISULFID	441
FT	DISULFID	457
FT	DISULFID	462
FT	DISULFID	479
FT	DISULFID	495
FT	DISULFID	500
FT	DISULFID	517
FT	DISULFID	533
FT	DISULFID	538
FT	DISULFID	555
FT	CARBOHYD	98
FT	CARBOHYD	137
FT	CARBOHYD	167
FT	CARBOHYD	649
FT	CONFLICT	437
FT	CONFLICT	459
FT	CONFLICT	490
SQ	SEQUENCE	880 AA; 94643 MW; E967E562 CRC32;
Query Match 98.1%; Score 6045; DB 1; Length 880;		
Best Local Similarity 98.0%; Pred. No. 0.00e+00;		
Matches 835; Conservative 9; Mismatches 7; Indels 1; Gaps 1;		
Db	1	MHWICLLTATCFVTVVQVHSSGSEFELRLKYFSNDHGRDNEGRCCSGESGATCKCLGS 60
Qy	1	MHWICLLTATCFVTVVQVHSSGSEFELRLKYFSNDHGRDNEGRCCSGESGATCKCLGS 60
Db	61	CKTRFCLCKHYQATIDTTSQCTYGDVITPILGENSVNLTDQRFQNGFTNPQPPFSF 120
Qy	61	CKTRFVCLCKHYQATIDTTSQCTYGDVITPILGENSVNLTDQRFQNGFTNPQPPFSF 120
Db	121	SWPGTFSLIVAEWHDNTNNGARTNKLTLQRLVQVLEVSSEWTKNSESQYTSLEYDF 180
Qy	121	SWPGTFSLIVAEWHDNTNNGARTNKLTLQRLVQVLEVSSEWTKNSESQYTSLEYDF 180
Db	181	RVTCDLNYGSGCAKFCRPRDSSFHSTCSETGEIICLTGWQGDYCHIPKCAKGCHEGHC 240
Qy	181	RVTCDLNYGSGCAKFCRPRDSSFHSTCSETGEIICLTGWQGDYCHIPKCAKGCHEGHC 240
Db	241	DKPNQCVQLGKWKALCNVLEPNCIHGTCKNKPWCICNEGGGLYCNQDLNYCTNHRP 300

Qy	241	DKPNQCVQLGKWKALCNVLEPNCIHGTCKNKPWCICNEGGGLYCNQDLNYCTNHRP 300
Db	301	CKNGTCTFNTGEGLYTCKCAPGYSGGDCENEIYSCADVNPQNGGTCTIDEPHRTGYKC 360
Qy	301	CKNGTCTFNTGEGLYTCKCAPGYSGGDCENEIYSCADVNPQNGGTCTIDEPHRTGYKC 360
Db	361	HCNWSGKMECEKVLCTCDKPCCHOGICRNVRPGLSGSGQGYQCEPCTGYSNCDLQLD 420
Qy	361	HCNWSGKMECEKVLCTCDKPCCHOGICRNVRPGLSGSGQGYQCEPCTGYSNCDLQLD 420
Db	421	NCSNPNCINGGSCQPSGKICPISGFSGRCTETINDDCLGHQCENGGTCTIDMWNQYRCQCV 480
Qy	421	NCSNPNCINGGSCQPSGKICPISGFSGRCTETINDDCLGHQCENGGTCTIDMWNQYRCQCV 480
Db	481	PGFHGTHGSSKVDLCLIRPCANGGTCLNNDYQCTCRAGFTGKDCSDVIDECSSGPCHN 540
Qy	481	PGFHGTHGSSKVDLCLIRPCANGGTCLNNDYQCTCRAGFTGKDCSDVIDECSSGPCHN 540
Db	541	GCTCMNRVNSPECVANGFRGKQDEESVTFDAHQYGATTQARADGLANAQVVLIAV 600
Qy	541	GCTCMNRVNSPECVANGFRGKQDEESVTFDAHQYGATTQARADGLANAQVVLIAV 600
Db	601	FSVAMPLVAVIAACVVFVCMKRRKRAQKDNAAERKQNEQNAVATMHNGSAGVALASA 660
Qy	601	FSVAMPLVAVIAACVVFVCMKRRKRAQKDNAAERKQNEQNAVATMHNGSAGVALASA 660
Db	661	SMGGTGSNSGLTFDGGNPNIKTWDSVNNICASAAAAAABCLMYGGYVASV 720
Qy	661	SMGGTGSNSGLTFDGGNPNIKTWDSVNNICASAAAAAABCLMYGGYVASV 720
Db	721	ADNNNANSDFCVAPLQRAKSOQLNTDPTLMHRGSPAGTSAGSGGGPAAEGKRISVL 780
Qy	721	ADNNNANSDFCVAPLQRAKSOQLNTDPTLMHRGSPAGTSAGSGGGPAAEGKRISVL 780
Db	781	GGGSCYCSQWPSLAAGVAGDLFIOLMAAASVAGTD-GTAQOQSVVCTGPH 831
Qy	781	GGGSCYCSQWPSLAAGVAGDLFIOLMAAASVAGTD-GTAQOQSVVCTGPH 831
RESULT 2		
ID	DLL1_MOUSE	STANDARD; PRT: 722 AA.
AC	O61483	
DT	01-NOV-1997	(REL. 35, CREATED)
DT	01-NOV-1997	(REL. 35, LAST SEQUENCE UPDATE)
DT	01-NOV-1997	(REL. 35, LAST ANNOTATION UPDATE)
DE	DELTA-LIKE PROTEIN 1 PRECURSOR (DELTA1).	
GN	DLL1.	
OS	MUS MUSCULUS (MOUSE).	
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;	
OC	EUTHERIA; RODENTIA.	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN-BALB/C X C57BL/6; TISSUE=EMBRYO;	
RX	MEDLINE; 95401858.	
RA	BETTENHAUSEN B., DE ANGELIS M.H., SIMON D., GUENET J.-L., GOSSLER A.;	
RL	DEVELOPMENT 121:2407-2418(1995).	
CC	-!- FUNCTION: MAY BE INVOLVED IN CELL-TO-CELL COMMUNICATION IN	
CC	MAMMALIAN EMBRYOS. MAY HAVE A ROLE IN CELLULAR INTERACTIONS	
CC	UNDERLYING SOMITOGENESIS AND DEVELOPMENT OF THE NERVOUS SYSTEM.	
CC	-!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.	
CC	-!- TISSUE SPECIFICITY: IN THE EMBRYO, EXPRESSED IN THE PARAXIAL	
CC	MESODERM AND NERVOUS SYSTEM. EXPRESSED AT HIGH LEVELS IN ADULT	
CC	HEART AND AT LOWER LEVELS, IN ADULT LUNG.	
CC	-!- DEVELOPMENTAL STAGE: EXPRESSED UNTIL DAY 15 IN THE EMBRYO.	
CC	-!- SIMILARITY: CONTAINS 8 EGF-LIKE DOMAINS.	
CC	-!- SIMILARITY: TO DROSOPHILA DELTA PROTEIN.	
DR	EMBL; X80903; G806570; -	
DR	MGI; MGI:104659; DLL1.	
DR	PROSITE; PS00010; ASX HYDROXYL; 3.	
DR	PROSITE; PS00022; EGF 1; 8.	
DR	PROSITE; PS01186; EGF 2; 8.	

DB	418	KCVLDGNSYLRCQCGFGRCENVDVDCASSPANGCSDSDSCPTCPPTGKNC	477
DB	418	KCVLDGNSYLRCQCGFGRCENVDVDCASSPANGCSDSDSCPTCPPTGKNC	477
OY	432	SCQSGK---CICPAGEGTGTCENIDDLCHQENSGTCLIDMANNQRCOCVPEFHGHC	488
DB	478	SAPVSRCEHAPCHNGATCQRCGORYMECEAGYGGGPNCOFLPPPPGP	526
OY	489	SSKVDLCILRPCANGCTCLINLNDYQCTCRAGFGKDCSDVIDECSSGP	537
RESULT	3	STANDARD;	PRT; 714 AA.
ID	DL1	RAT	
AC	P97677		
DT	01-NOV-1997	(REL. 35, CREATED)	
DT	01-NOV-1997	(REL. 35, LAST SEQUENCE UPDATE)	
DT	01-NOV-1997	(REL. 35, LAST ANNOTATION UPDATE)	
DE	DELTA-LIKE PROTEIN 1 PRECURSOR (DELTA1).		
GN	DL1.		
OS	RATTUS NORVEGICUS (RAT).		
OC	EUMAROTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;		
OC	EUTHERIA; RODENTIA.		
RN	(1)		
RP	SEQUENCE FROM N.A.		
RA	DISIBIO G., HEBISHI L., BOULTER J., WEINMASTER G.;		
RL	SUBMITTED (DEC-1996) TO EMBL/GENBANK/DBJ DATA BANKS.		
CC	-1- FUNCTION: MAY BE INVOLVED IN CELL-TO-CELL COMMUNICATION IN		
CC	MAMMALIAN EMBRYOS. MAY HAVE A ROLE IN CELLULAR INTERACTIONS		
CC	UNDERLYING SOMITOGENESIS AND DEVELOPMENT OF THE NERVOUS SYSTEM (H		
CC	SMILARITY).		
CC	-1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.		
CC	-1- SIMILARITY: CONTAINS 8 EGF-LIKE DOMAINS.		
CC	-1- SIMILARITY: TO DROSOPHILA DELTA PROTEIN.		
DR	EMBL, U78889; G1699046; --		
DR	PROSITE; PS00010; ASX_HYDROXYL; 3.		
DR	PROSITE; PS00022; EGF_1; 8.		
DR	PROSITE; PS01186; EGF_2; 8.		
DR	PROSITE; PS01187; EGF_CA; 2.		
KW	SIGNAL: EGF-LIKE DOMAIN; GLYCOPROTEIN; TRANSMEMBRANE.		
FT	SIGNAL	1	17
FT	CHAIN	18	714
FT	DOMAIN	18	537
FT	TRANSMEM	538	560
FT	DOMAIN	561	714
FT	DOMAIN	225	253
FT	DOMAIN	256	284
FT	DOMAIN	291	324
FT	DOMAIN	331	362
FT	DOMAIN	369	401
FT	DOMAIN	408	439
FT	DOMAIN	446	477
FT	DOMAIN	484	515
FT	DOMAIN	225	236
FT	DISULFID	229	242
FT	DISULFID	244	253
FT	DISULFID	256	267
FT	DISULFID	267	273
FT	DISULFID	275	284
FT	DISULFID	291	303
FT	DISULFID	297	313
FT	DISULFID	315	324
FT	DISULFID	331	342
FT	DISULFID	336	351
FT	DISULFID	353	362
FT	DISULFID	369	380
FT	DISULFID	374	390
FT	DISULFID	392	401
FT	DISULFID	408	419
FT	DISULFID	413	428
FT	DISULFID	430	439
FT	DISULFID	446	466
FT	DISULFID	468	477
FT	DISULFID	484	495
FT	DISULFID	489	504
FT	DISULFID	506	515

OF EMBRYO IS ONE OF THE FIRST STEPS OF CNS DEVELOPMENT IN INSECTS.
THIS PROCESS IS UNDER CONTROL OF THE NEUROGENIC GENES.
-!- NOTCH AND SERRATE MAY INTERACT AT THE PROTEIN LEVEL, IT IS CONCEIVABLE THAT THE SERRATE AND DELTA PROTEINS MAY COMPETE FOR BINDING WITH THE NOTCH PROTEIN.
-!- SIMILARITY: CONTAINS 14 EGF-LIKE DOMAINS.
-!- SIMILARITY: TO THE DROSOPHILA NEUROGENIC LOCUS DELTA PROTEIN.
EMBL: X56811; G8564; -.
EMBL: M35759; G158606; -.
PIR: A36666; A36666.
PIR: S16878; S16878.
HSP: P00743; IAP0.
FLYBASE: FBgn0004197; Ser.
PROSITE: PS00010; ASX_HYDROXYL; 7.
PROSITE: PS00022; EGF_1; 14.
PROSITE: PS01186; EGF_2; 8.
PROSITE: PS01187; EGF_CA; 5.
DIFFERENTIATION; REPEAT; EGF-LIKE DOMAIN; TRANSMEMBRANE;
GLYCOPROTEIN; SIGNAL.
FT CHAIN 1 83 POTENTIAL.
FT CHAIN 84 1408 SERRATE PROTEIN.
FT DOMAIN 84 1223 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1224 1249 POTENTIAL.
FT DOMAIN 1250 1408 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 284 317 EGF-LIKE 1.
FT DOMAIN 315 349 EGF-LIKE 2.
FT DOMAIN 351 389 EGF-LIKE 3.
FT DOMAIN 391 489 EGF-LIKE 4.
FT DOMAIN 407 476 SER-RICH (INSERT).
FT DOMAIN 491 527 EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 529 609 EGF-LIKE 6, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 611 646 EGF-LIKE 7, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 648 684 EGF-LIKE 8, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 686 721 EGF-LIKE 9.
FT DOMAIN 723 797 EGF-LIKE 10.
FT DOMAIN 737 769 THR-RICH (INSERT).
FT DOMAIN 799 835 EGF-LIKE 11, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 837 877 EGF-LIKE 12.
FT DOMAIN 879 915 EGF-LIKE 13.
FT DOMAIN 917 953 EGF-LIKE 14, CALCIUM-BINDING (POTENTIAL).
FT DISULFID 288 299 BY SIMILARITY.
FT DISULFID 292 305 BY SIMILARITY.
FT DISULFID 307 316 BY SIMILARITY.
FT DISULFID 319 330 BY SIMILARITY.
FT DISULFID 325 337 BY SIMILARITY.
FT DISULFID 339 348 BY SIMILARITY.
FT DISULFID 355 367 BY SIMILARITY.
FT DISULFID 369 377 BY SIMILARITY.
FT DISULFID 379 388 BY SIMILARITY.
FT DISULFID 395 406 BY SIMILARITY.
FT DISULFID 400 477 BY SIMILARITY.
FT DISULFID 479 488 BY SIMILARITY.
FT DISULFID 495 506 BY SIMILARITY.
FT DISULFID 500 515 BY SIMILARITY.
FT DISULFID 517 526 BY SIMILARITY.
FT DISULFID 533 588 BY SIMILARITY.
FT DISULFID 582 597 BY SIMILARITY.
FT DISULFID 599 608 BY SIMILARITY.
FT DISULFID 615 625 BY SIMILARITY.
FT DISULFID 619 634 BY SIMILARITY.
FT DISULFID 636 645 BY SIMILARITY.
FT DISULFID 652 663 BY SIMILARITY.
FT DISULFID 657 672 BY SIMILARITY.
FT DISULFID 674 683 BY SIMILARITY.
FT DISULFID 690 700 BY SIMILARITY.
FT DISULFID 695 709 BY SIMILARITY.
FT DISULFID 711 720 BY SIMILARITY.
FT DISULFID 803 814 BY SIMILARITY.
FT DISULFID 808 823 BY SIMILARITY.
FT DISULFID 825 834 BY SIMILARITY.
FT DISULFID 841 852 BY SIMILARITY.
FT DISULFID 846 865 BY SIMILARITY.
FT DISULFID 867 876 BY SIMILARITY.

[illegible]

DT		01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)	
DE		01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)	
DI		NEUROGENIC LOCUS NOTCH HOMOLOG PROTEIN PRECURSOR.	
DN		NOTCH.	
GN		BRACHYDANTIO RERIO (ZEBRAFISH) (ZEBRA DANTIO).	
OC	EUKARYOTA; METAFAA; CHORODTA; VERTEBRATA; PICES;	GNAHOSTOMATA.	
OC	OSEFICHTHES; ACTINOPTERYGII; CYPRINIFORMES.		
RN	[1]	SEQUENCE FROM N.A.	
RC		TISSUE=EMBRYO;	
RX	MEDLINE: 94128602.		
RA	BIEKKAP C.; CAMPOS-ORTEGA J.A.;		
RL	MECH. DEV. 43:87-100(1993).		
CC	-I-	FUNCTION: IMPLICATED IN CELL FATE SPECIFICATIONS DURING EMBRIO DEVELOPMENT. MAY BE INVOLVED IN THE FORMATION OF THE NEURAL PLATE, NOTCHORD AND BRAIN VESICLES.	
CC	-I-	SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.	
CC	-I-	DEVELOPMENTAL STAGE: EXPRESSED IN ALL CELLS IN PRGASURULATION STAGES. DURING GASTRULATION IS DIFFERENTIALLY EXPRESSED ACCUMULATING PREDOMINANTLY IN THE PRECHORDAL MESODERM AND NOTOCHORD. AT THE END OF GASTRULATION, EXPRESSED ALONG THE ANTERIOR-POSTERIOR AXIS INCLUDING THE DEVELOPING NEURAL PLATE AND DIFFERENTIATING MESODERM. ALSO PRESENT IN THE DEVELOPING BRAIN AND HEAD REGIONS.	
CC	-I-	SIMILARITY: HIGH, WITH OTHER NOTCH-TYPE PROTEINS.	
CC	-I-	SIMILARITY: CONTAINS 36 EGF-LIKE DOMAINS.	
CC	-I-	SIMILARITY: THIS PROTEIN INCLUDES 3 COPIES OF A LIN/NOTCH REPEAT.	
CC	-I-	SIMILARITY: CONTAINS 6 ANK REPEATS.	
DR	EMBL: X69088; G433867; -.		
DR	PROSITE: PS00010; ASX_HYDROXL; 23.		
DR	PROSITE: PS00022; EGF_1; 34.		
DR	PROSITE: PS01186; EGF_2; 28.		
DR	PROSITE: PS01187; EGF_CA; 22.		
KW	DIFFERENTIATION; NEUROGENESIS; REPEAT; ANK REPEAT; EGF-LIKE DOMAIN; TRANSMEMBRANE; SIGNAL; GLYCOPROTEIN.		
FT	SIGNAL	1	20
FT	CHAIN	21	2437
FT	DOMAIN	21	1724
FT	TRANSMEM	1725	1747
FT	DOMAIN	1748	2437
FT	DOMAIN	21	57
FT	DOMAIN	58	98
FT	DOMAIN	101	138
FT	DOMAIN	139	175
FT	DOMAIN	177	215
FT	DOMAIN	217	254
FT	DOMAIN	256	292
FT	DOMAIN	294	332
FT	DOMAIN	334	370
FT	DOMAIN	371	409
FT	DOMAIN	411	449
FT	DOMAIN	451	487
FT	DOMAIN	489	524
FT	DOMAIN	526	562
FT	DOMAIN	564	599
FT	DOMAIN	601	637
FT	DOMAIN	639	674
FT	DOMAIN	676	712
FT	DOMAIN	714	749
FT	DOMAIN	751	787
FT	DOMAIN	789	825
FT	DOMAIN	827	865
FT	DOMAIN	867	903
FT	DOMAIN	905	941
FT	DOMAIN	943	979
FT	DOMAIN	981	1017
FT	DOMAIN	1019	1055
FT	DOMAIN	1057	1093
FT	DOMAIN	1095	1141
FT	DOMAIN	1143	1179
FT	DOMAIN	1181	1217
FT	DOMAIN	1219	1263
FT	DOMAIN	1265	1303

```
FT DOMAIN 1305 1344 EGF-LIKE 34. 775 760 775 BY SIMILARITY.
FT DOMAIN 1346 1382 EGF-LIKE 35. 786 777 786 BY SIMILARITY.
FT DOMAIN 1385 1423 EGF-LIKE 36. 804 793 804 BY SIMILARITY.
FT DOMAIN 1446 1486 3 X LIN/NOTCH REPEATS. 813 798 813 BY SIMILARITY.
FT REPEAT 1487 1520 LIN/NOTCH 1. 824 815 824 BY SIMILARITY.
FT REPEAT 1487 1520 LIN/NOTCH 2. 831 842 831 BY SIMILARITY.
FT REPEAT 1521 1561 LIN/NOTCH 3. 853 836 853 BY SIMILARITY.
FT DOMAIN 1861 2074 6 X ANK MOTIF REPEATS. 864 855 864 BY SIMILARITY.
FT REPEAT 1861 1891 ANK MOTIF 1. 882 871 882 BY SIMILARITY.
FT REPEAT 1892 1940 ANK MOTIF 1. 891 876 891 BY SIMILARITY.
FT REPEAT 1941 1974 ANK MOTIF 1. 902 893 902 BY SIMILARITY.
FT REPEAT 1975 2007 ANK MOTIF 1. 929 909 929 BY SIMILARITY.
FT REPEAT 2008 2040 ANK MOTIF 1. 940 914 940 BY SIMILARITY.
FT REPEAT 2041 2074 ANK MOTIF 1. 958 931 958 BY SIMILARITY.
FT DOMAIN 2265 2276 POLY-GLN (OPA-REPEAT). 967 947 967 BY SIMILARITY.
FT DOMAIN 25 35 BY SIMILARITY. 978 969 978 BY SIMILARITY.
FT DISULFID 29 45 BY SIMILARITY. 1034 1023 1034 BY SIMILARITY.
FT DISULFID 47 56 BY SIMILARITY. 1043 1028 1043 BY SIMILARITY.
FT DISULFID 62 73 BY SIMILARITY. 1054 1045 1054 BY SIMILARITY.
FT DISULFID 67 86 BY SIMILARITY. 1072 1061 1072 BY SIMILARITY.
FT DISULFID 88 97 BY SIMILARITY. 1081 1066 1081 BY SIMILARITY.
FT DISULFID 105 116 BY SIMILARITY. 1092 1083 1092 BY SIMILARITY.
FT DISULFID 110 126 BY SIMILARITY. 1120 1099 1120 BY SIMILARITY.
FT DISULFID 128 137 BY SIMILARITY. 1129 1114 1129 BY SIMILARITY.
FT DISULFID 143 154 BY SIMILARITY. 1140 1131 1140 BY SIMILARITY.
FT DISULFID 148 163 BY SIMILARITY. 1158 1147 1158 BY SIMILARITY.
FT DISULFID 165 174 BY SIMILARITY. 1167 1152 1167 BY SIMILARITY.
FT DISULFID 181 194 BY SIMILARITY. 1178 1169 1178 BY SIMILARITY.
FT DISULFID 188 203 BY SIMILARITY. 1196 1185 1196 BY SIMILARITY.
FT DISULFID 205 214 BY SIMILARITY. 1205 1190 1205 BY SIMILARITY.
FT DISULFID 221 232 BY SIMILARITY. 1216 1207 1216 BY SIMILARITY.
FT DISULFID 226 242 BY SIMILARITY. 1242 1223 1242 BY SIMILARITY.
FT DISULFID 244 253 BY SIMILARITY. 1251 1236 1251 BY SIMILARITY.
FT DISULFID 260 271 BY SIMILARITY. 1262 1253 1262 BY SIMILARITY.
FT DISULFID 265 280 BY SIMILARITY. 1282 1269 1282 BY SIMILARITY.
FT DISULFID 282 291 BY SIMILARITY. 1291 1274 1291 BY SIMILARITY.
FT DISULFID 298 311 BY SIMILARITY. 1302 1293 1302 BY SIMILARITY.
FT DISULFID 305 320 BY SIMILARITY. 1320 1309 1320 BY SIMILARITY.
FT DISULFID 322 331 BY SIMILARITY. 1332 1314 1332 BY SIMILARITY.
FT DISULFID 338 349 BY SIMILARITY. 1343 1334 1343 BY SIMILARITY.
FT DISULFID 343 358 BY SIMILARITY. 1361 1350 1361 BY SIMILARITY.
FT DISULFID 360 369 BY SIMILARITY. 1370 1355 1370 BY SIMILARITY.
FT DISULFID 375 386 BY SIMILARITY. 1381 1372 1381 BY SIMILARITY.
FT DISULFID 380 397 BY SIMILARITY. 1400 1389 1400 BY SIMILARITY.
FT DISULFID 399 408 BY SIMILARITY. 1411 1394 1411 BY SIMILARITY.
FT DISULFID 415 428 BY SIMILARITY. 1422 1413 1422 BY SIMILARITY.
FT DISULFID 432 443 BY SIMILARITY. 957 957 POTENTIAL.
FT DISULFID 439 448 BY SIMILARITY. 1177 1177 POTENTIAL.
FT DISULFID 455 466 BY SIMILARITY. 1487 1487 POTENTIAL.
FT DISULFID 460 475 BY SIMILARITY. 1585 1585 POTENTIAL.
FT DISULFID 477 486 BY SIMILARITY.
FT DISULFID 493 503 BY SIMILARITY.
FT DISULFID 512 523 BY SIMILARITY.
FT DISULFID 514 523 BY SIMILARITY.
FT DISULFID 530 541 BY SIMILARITY.
FT DISULFID 535 550 BY SIMILARITY.
FT DISULFID 552 561 BY SIMILARITY.
FT DISULFID 568 578 BY SIMILARITY.
FT DISULFID 573 587 BY SIMILARITY.
FT DISULFID 589 598 BY SIMILARITY.
FT DISULFID 605 616 BY SIMILARITY.
FT DISULFID 610 625 BY SIMILARITY.
FT DISULFID 627 636 BY SIMILARITY.
FT DISULFID 643 653 BY SIMILARITY.
FT DISULFID 648 662 BY SIMILARITY.
FT DISULFID 664 673 BY SIMILARITY.
FT DISULFID 680 691 BY SIMILARITY.
FT DISULFID 685 700 BY SIMILARITY.
FT DISULFID 702 711 BY SIMILARITY.
FT DISULFID 718 728 BY SIMILARITY.
FT DISULFID 723 737 BY SIMILARITY.
FT DISULFID 739 748 BY SIMILARITY.
FT DISULFID 755 766 BY SIMILARITY.
```

... Note: remainder of annotations omitted.

Query Match 16.08; Score 987; DB 1; Length 2437;
Best Local Similarity 39.18; Pred. No. 1.98e-212; Mismatches 124; Indels 34; Gaps 23;
Matches 154; Conservative 82;

```
Db 643 CKRRPCDYGKIDKINGEYCEVCEPGYSGSMCNINIDCALNPGHNGTICDGVNSFTCLC 702
QY 197 CRPRDSDFGHSTCSETG-EIICLTGWQGDYCHIP--KCA-KGCEHG-HC-DKPNQ--CVC 248
Db 703 PDGFRDATCISQHNCCSNP-CHGSCLDQINSYRCVCEAGWGRNCDININECLSN-PC 760
QY 249 QLGWKGALC----NECVLEPNCIHGTC-N--KPWTCICNEGWGLYCNQDLNYCTNHRPC 301
Db 761 VNGGTCKDMTSG-YLCTCRAGFSFGSPNCQMINEC-AS-NPCLNCGSCIDD--V-AGEKCN 814
QY 302 KNGGTCFNTGEGLYTCRCAPYSGDCCENIEYISDADVNPCQNGGTICIDSPHTKTGYKCH 361
Db 815 CMLPYTCEVCEN-VLAPCSPRCPCKNGGVCRE-SEDF----QSFSNCNCPAGWQGTCEVDI 868
QY 362 CANGWSKMKCEKVLTC-CDKPCHQG-ICRNVPRGLSGKGGYQCECPIGYSGPNCIDLQ 419
```


FT	DISULFID	150	165	BY SIMILARITY.
FT	DISULFID	167	176	BY SIMILARITY.
FT	DISULFID	183	194	BY SIMILARITY.
FT	DISULFID	188	203	BY SIMILARITY.
FT	DISULFID	205	214	BY SIMILARITY.
FT	DISULFID	221	232	BY SIMILARITY.
FT	DISULFID	226	242	BY SIMILARITY.
FT	DISULFID	244	253	BY SIMILARITY.
FT	DISULFID	260	271	BY SIMILARITY.
FT	DISULFID	265	280	BY SIMILARITY.
FT	DISULFID	282	291	BY SIMILARITY.
FT	DISULFID	298	311	BY SIMILARITY.
FT	DISULFID	305	320	BY SIMILARITY.
FT	DISULFID	322	331	BY SIMILARITY.
FT	DISULFID	338	349	BY SIMILARITY.
FT	DISULFID	343	358	BY SIMILARITY.
FT	DISULFID	360	369	BY SIMILARITY.
FT	DISULFID	375	386	BY SIMILARITY.
FT	DISULFID	380	397	BY SIMILARITY.
FT	DISULFID	399	408	BY SIMILARITY.
FT	DISULFID	415	428	BY SIMILARITY.
FT	DISULFID	422	437	BY SIMILARITY.
FT	DISULFID	439	448	BY SIMILARITY.
FT	DISULFID	455	466	BY SIMILARITY.
FT	DISULFID	460	475	BY SIMILARITY.
FT	DISULFID	477	486	BY SIMILARITY.
FT	DISULFID	493	504	BY SIMILARITY.
FT	DISULFID	498	513	BY SIMILARITY.
FT	DISULFID	515	524	BY SIMILARITY.
FT	DISULFID	531	542	BY SIMILARITY.
FT	DISULFID	536	551	BY SIMILARITY.
FT	DISULFID	553	562	BY SIMILARITY.
FT	DISULFID	569	579	BY SIMILARITY.
FT	DISULFID	574	588	BY SIMILARITY.
FT	DISULFID	590	599	BY SIMILARITY.
FT	DISULFID	606	617	BY SIMILARITY.
FT	DISULFID	611	626	BY SIMILARITY.
FT	DISULFID	628	637	BY SIMILARITY.
FT	DISULFID	644	654	BY SIMILARITY.
FT	DISULFID	649	663	BY SIMILARITY.
FT	DISULFID	665	674	BY SIMILARITY.
FT	DISULFID	681	692	BY SIMILARITY.
FT	DISULFID	686	701	BY SIMILARITY.
FT	DISULFID	703	712	BY SIMILARITY.
FT	DISULFID	719	729	BY SIMILARITY.
FT	DISULFID	724	738	BY SIMILARITY.
FT	DISULFID	740	749	BY SIMILARITY.
FT	DISULFID	756	767	BY SIMILARITY.
FT	DISULFID	761	776	BY SIMILARITY.
FT	DISULFID	778	787	BY SIMILARITY.
FT	DISULFID	794	805	BY SIMILARITY.
FT	DISULFID	799	814	BY SIMILARITY.
FT	DISULFID	816	825	BY SIMILARITY.
FT	DISULFID	832	843	BY SIMILARITY.
FT	DISULFID	837	854	BY SIMILARITY.
FT	DISULFID	856	865	BY SIMILARITY.
FT	DISULFID	872	883	BY SIMILARITY.
FT	DISULFID	877	892	BY SIMILARITY.
FT	DISULFID	894	903	BY SIMILARITY.
FT	DISULFID	910	921	BY SIMILARITY.
FT	DISULFID	915	930	BY SIMILARITY.
FT	DISULFID	932	941	BY SIMILARITY.
FT	DISULFID	986	997	BY SIMILARITY.
FT	DISULFID	991	1006	BY SIMILARITY.
FT	DISULFID	1008	1017	BY SIMILARITY.
FT	DISULFID	1024	1035	BY SIMILARITY.
FT	DISULFID	1029	1044	BY SIMILARITY.
FT	DISULFID	1046	1055	BY SIMILARITY.
FT	DISULFID	1062	1073	BY SIMILARITY.
FT	DISULFID	1067	1082	BY SIMILARITY.
FT	DISULFID	1084	1093	BY SIMILARITY.
FT	DISULFID	1100	1121	BY SIMILARITY.
FT	DISULFID	1115	1130	BY SIMILARITY.

FT	DISULFID	1132	1141	BY SIMILARITY.
FT	DISULFID	1148	1159	BY SIMILARITY.
FT	DISULFID	1153	1168	BY SIMILARITY.
FT	DISULFID	1170	1179	BY SIMILARITY.
FT	DISULFID	1186	1197	BY SIMILARITY.
FT	DISULFID	1191	1206	BY SIMILARITY.
FT	DISULFID	1208	1217	BY SIMILARITY.
FT	DISULFID	1224	1243	BY SIMILARITY.
FT	DISULFID	1237	1252	BY SIMILARITY.
FT	DISULFID	1254	1263	BY SIMILARITY.
FT	DISULFID	1270	1283	BY SIMILARITY.
FT	DISULFID	1275	1292	BY SIMILARITY.
FT	DISULFID	1294	1303	BY SIMILARITY.
FT	DISULFID	1310	1321	BY SIMILARITY.
FT	DISULFID	1315	1333	BY SIMILARITY.
FT	DISULFID	1335	1344	BY SIMILARITY.
FT	DISULFID	1351	1362	BY SIMILARITY.
FT	DISULFID	1356	1371	BY SIMILARITY.
FT	DISULFID	1373	1382	BY SIMILARITY.
FT	DISULFID	1390	1401	BY SIMILARITY.
FT	DISULFID	1395	1412	BY SIMILARITY.
FT	DISULFID	1414	1423	BY SIMILARITY.
FT	DISULFID	1414	1423	POTENTIAL.
FT	CARBOHYD	462	462	POTENTIAL.
FT	CARBOHYD	887	887	POTENTIAL.
FT	CARBOHYD	958	958	POTENTIAL.
FT	CARBOHYD	1178	1178	POTENTIAL.
FT	CARBOHYD	1487	1487	POTENTIAL.
FT	CARBOHYD	1508	1508	POTENTIAL.
FT	CARBOHYD	1584	1584	POTENTIAL.
SEQ	SEQUENCE	2524 AA;	275123 MW;	E63584B1 CRC32;

Query Match 15.5%; Score 956; DB 1; Length 2524;

Best Local Similarity 36.0%; Pred. No. 2.01e-204; Matches 134; Conservative 96; Mismatches 112; Indels 30; Gaps 24;

Db	221	CNPSPLNGGTORQD-DTSYDCTCLPGRSGNCEENIDCPSNCRNGTCVDGVNFTN	279
Qy	226	CHPRKARK--CEHGCDPRNOCYQOLGKWKALCNCEVLE-P--KNIHG-TC-N--KPMT	276
Db	280	COCPPDMTQYCTEDVDECOLMPNACNGGTCHNTYGG--YNCVYNGWTGEDCSENIDDC	338
Qy	277	CICNEGWMGLYCNODLNYCT-NHRPCKNGGTCHNTGEGLYTCKAGYSGDCCENETYS	335
Db	339	-ANA-ACHSGARC-HD-RVASFY-CECPHGRGTLCHLDN-ACISNPKCEGSGNC-DTNP	390
Qy	336	DADVNCONGGTCTDEPHRTGTGKCHCANGWSKCEKRYLCTSDRPHQGI-CRVVRG	394
Db	391	VNGKA--I--CTCPPGYTGPACNNVDGCSIGANPCHEGRCCTWTLSFGQCNPOGYAGR	447
Qy	395	LQSKGGGYCCBPPIGYSGNCDLQDNGS--PWPCTNGSCOPS-G--KCIAPAGSGR	449
Db	448	CEIDVNECLSNPCONDSTGLDQIGEFQICIMPEEGLYCEINIDECASNPCLHNGKCIDK	507
Qy	450	CEINIDDCGHGHCENGTIDMVNQRCCVGFHGTSSKVDLLIRPCANGGCTNL	509
Db	508	INEFRCDPTGSGNLQCHDEDECTSTPCKNGAKCIDGNSNYTCCOCTBEFTGRHCEQDIN	567
Qy	510	NNDYOCCTCRAGFTGRDCSDVIDECSSGPHNGGTCNNRVNSFEVCANGFRGKOCCEES	569
Db	568	ECIP-DPCHYGT	578
Qy	570	DSVTFDAHOYGA	581
RESULT	8	STANDARD:	PRT: 2703 AA.
ID	NOTC-DROME		
AC	P07207; P04154;		
DT	01-NOV-1986 (REL. 03, CREATED)		
DT	01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)		
DT	01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)		
DE	NEUROGENIC LOCUS NOTCH PROTEIN PRECURSOR.		
GN	N.		
-OS	DROSOPHILA MELANOGASTER (FRUIT FLY).		

FT DISULFID 590 599 BY SIMILARITY.
FT DISULFID 606 616 BY SIMILARITY.
FT DISULFID 611 625 BY SIMILARITY.
FT DISULFID 627 636 BY SIMILARITY.
FT DISULFID 643 654 BY SIMILARITY.
FT DISULFID 648 663 BY SIMILARITY.
FT DISULFID 665 674 BY SIMILARITY.
FT DISULFID 681 692 BY SIMILARITY.
FT DISULFID 686 701 BY SIMILARITY.
FT DISULFID 703 712 BY SIMILARITY.
FT DISULFID 719 730 BY SIMILARITY.
FT DISULFID 724 739 BY SIMILARITY.
FT DISULFID 741 750 BY SIMILARITY.
FT DISULFID 757 768 BY SIMILARITY.
FT DISULFID 762 777 BY SIMILARITY.
FT DISULFID 779 788 BY SIMILARITY.
FT DISULFID 795 806 BY SIMILARITY.
FT DISULFID 800 815 BY SIMILARITY.
FT DISULFID 817 826 BY SIMILARITY.
FT DISULFID 833 844 BY SIMILARITY.
FT DISULFID 838 853 BY SIMILARITY.
FT DISULFID 855 864 BY SIMILARITY.
FT DISULFID 871 882 BY SIMILARITY.
FT DISULFID 876 893 BY SIMILARITY.
FT DISULFID 895 904 BY SIMILARITY.
FT DISULFID 911 923 BY SIMILARITY.
FT DISULFID 917 932 BY SIMILARITY.
FT DISULFID 934 943 BY SIMILARITY.
FT DISULFID 950 961 BY SIMILARITY.
FT DISULFID 955 970 BY SIMILARITY.
FT DISULFID 972 981 BY SIMILARITY.
FT DISULFID 988 999 BY SIMILARITY.
FT DISULFID 993 1008 BY SIMILARITY.
FT DISULFID 1010 1019 BY SIMILARITY.
FT DISULFID 1026 1037 BY SIMILARITY.
FT DISULFID 1031 1046 BY SIMILARITY.
FT DISULFID 1048 1057 BY SIMILARITY.
FT DISULFID 1064 1075 BY SIMILARITY.
FT DISULFID 1069 1084 BY SIMILARITY.
FT DISULFID 1086 1095 BY SIMILARITY.
FT DISULFID 1102 1113 BY SIMILARITY.
FT DISULFID 1107 1122 BY SIMILARITY.
FT DISULFID 1124 1133 BY SIMILARITY.
FT DISULFID 1140 1160 BY SIMILARITY.
FT DISULFID 1155 1169 BY SIMILARITY.
FT DISULFID 1171 1180 BY SIMILARITY.

Note: remainder of annotations omitted.

Query Match 15.4%; Score 951; DB 1; Length 2703;

Best Local Similarity 39.1%; Pred. No. 3, 91e-203; Matches 148; Conservative 83; Mismatches 116; Indels 32; Gaps 24;

Db 488 NINCESHPQNEGSCLDPEFTFRVCMPTGTCCELDIDECOSNPLNOSTCHDKING 547
Qy 227 HIPRC-ANGCH-GHC-DKPN--QCVCOLGWMKALCN-ECVL-EPN-GIH-GTC-NK--P 274
Db 548 FKSCALGFTGARCOINIDDCOSQ-PCRNRGICHDISIAG-YSCCEPPGYTGTSCININD 605
Qy 275 WTCICNEBMSGLYCNODLNTNHRPCKNRGTCFTGSLYTKCARPYSGDCENEITS 334
Db 606 CDS--NPHRG-KCID--VNS-FKCLDPEYTYICQKQINECSNPFQEDGHOD-R- 657
Qy 335 CDADVNPQNGGTCIDEPHTKTYGKCHCANGSMKCEKVLTCSDKPKCH--QGICRNVRP 393
Db 658 -VGS-----YYCQCAAGTSGKNCVAVNECHSNPCNMGATCIDGINSYKCCVPGTSGQHC 712
Qy 394 GLGSKGQGYCECPIGSGPNCIDLQDMCSNPGCINGSC-OP-SG-KCICPAGSGSTRC 450
Db 713 EKNVDECISSPCANNNGVYIDQVNGKCEPGEYDAHCTLDVDECASNPCVNEGRCEGOI 772
Qy 451 ETNIDDCILGHOCENGCTCIDMNVNTRCQCVPEFHGTHBSSKVDLCLIRPCANGTCLNLN 510

Db 773 NEFICHPGYTGRKCELDIDECSSNPGCHGCTCYDKLNAFSCQCMPTGTGKCEINID 832
Qy 511 NDYOCTCRAGFTGKDCSDVIDIDECSSPGCHNGCTCMNRVNSFECVCANGRGQCEDESD 570
Db 833 CVT-NPCNGGTCIDKRVNG 850
Qy 571 SVTFDAHGYGATQARADG 589
RESULT 9
ID NTCL_MOUSE STANDARD; PRI; 2531 AA.
AC Q01705;
DE 01-NOV-1995 (REL. 32, CREATED)
DE 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
DE 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
DE NEUROGENIC LOCUS NOTCH HOMOLOG PROTEIN 1 PRECURSOR (NOTCH PROTEIN).
GN NOTCH1 OR NOTCH.
OS MUS MUSCULUS (MOUSE).
OC EURARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RN SEQUENCE FROM N.A.
RP TISSUE-EMBRYO;
RX MEDLINE: 93194170.
RA FRANCO DEL AMO F., GENDRON-MAGUIRE M., SWIATEK P.J., JENKINS N.A.,
RA COPELAND N.G., GRIDLEY T.;
RN GENOMICS 15:259-264(1993).
[2]
RP SEQUENCE OF 1551-2170 FROM N.A.
RC TISSUE-EMBRYO;
RX MEDLINE: 93048835.
RA FRANCO DEL AMO F., SMITH D.E., SWIATEK P.J., GENDRON-MAGUIRE M.,
RA GREENSPAN R.J., MCMAHON A.P., GRIDLEY T.;
RL DEVELOPMENT 115:737-744(1992).
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED ALMOST UNIFORMLY IN EARLY EMBRYOS.
CC -1- SIMILARITY: CONTAINS 36 EGF-LIKE DOMAINS.
CC -1- SIMILARITY: THIS PROTEIN INCLUDES 3 COPIES OF A LIN/NOTCH REPEAT.
CC -1- SIMILARITY: CONTRAINS 6 ANK REPEATS.
CC -1- SIMILARITY: HIGH, WITH OTHER NOTCH-TYPE PROTEINS.
DR EMBL: 211886; G288503; -
DR MGI: 97363; NOTCH1.
DR PROSITE: PS0010; ASX_HYDROXYL; 22.
DR PROSITE: PS00022; EGF_1; 34.
DR PROSITE: PS01186; EGF_2; 27.
DR PROSITE: PS01187; EGF_CA; 21.
DR DIFFERENTIATION; NEUROGENESIS; REPEAT; ANK REPEAT; EGF-LIKE DOMAIN;
KW TRANSMEMBRANE; SIGNAL; GLYCOPROTEIN.
FT SIGNAL 1 18
FT CHAIN 19 2531
FT DOMAIN 19 1725
FT TRANSMEM 1726 1746
FT DOMAIN 1747 2531
FT DOMAIN 24 1425
FT DOMAIN 1449 1462
FT DOMAIN 1445 1562
FT DOMAIN 1445 1480
FT REPEAT 1481 1522
FT REPEAT 1523 1562
FT DOMAIN 1865 2075
FT REPEAT 1865 1910
FT REPEAT 1912 1942
FT REPEAT 1944 1975
FT REPEAT 1978 2009
FT REPEAT 2011 2042
FT REPEAT 2044 2075
FT CARBOHYD 888 888
FT CARBOHYD 959 959
FT CARBOHYD 1179 1179
FT CARBOHYD 1241 1241
FT CARBOHYD 1489 1489
FT CARBOHYD 1587 1587
FT SEQUENCE 2531 AA; 271312 MW; AD71189B CRC32;

Query Match	15.1%	Score 928;	DB 1;	Length 2531;
Best Local Similarity	38.7%	Pred. No. 3.32e-197;		
Matches	147;	Conservative	82;	Mismatches 113; Indels 38; Gaps 26;
Db	465	ATCLDQIGFQCICMPGYGVYCEINTEDEACSSPCLNHGCMKIHIEFQCQCPKFGNHL	524	
Qy	207	STC-SETGEI--ICLTGWGDYCHI-PK-CAKG-CHEH-DK-PN-QCVCQLGKMGAL	256	
Db	525	COYDVDECASTP-CCKNAKCLDGPNTVCTCYGTTCHCEVIDEC-DPDPCHYG-SC-	580	
Qy	257	C----NECVLEPNCHGT-C-NKP--WTCINCEGWGLYCNQDLNYCTNHRPCNKNGTCF	308	
Db	581	KDGVATETCLCQCPQYGHGHTETNIECHSQ--PCRHGTC-QD-RDMS-YLCICLKRGTTG	635	
Qy	309	NTGBGLYTCAPGYSGDDCENEIYSCDADVNPQNGGTCIDEPHTKTGYKCHCANGWG	368	
Db	636	PNCBINDDCASPCDSGTC-----LD-KIDGYECACEPGYGSGMCNVNIDECASPC	688	
Qy	369	KMCEKVLTCSDKPGCHOGICRNVPRPLGSKGGYQCPCPIGYSGPNGLDLQDNCSPNCI	428	
Db	689	NGGFCEDGIAGFTCRCPGYYDPTCLSEVNECNSNPCHG-ACRDGLNGYKCDACGWSG	747	
Qy	429	NGGSCQP--SG-KCICPAGFSGTCTETNIDCLGHQCEGNGTCTDMVNYQRCVCPFGH	485	
Db	748	TNCDINNECESPCVANGGTCOMTSGYVCTCREGFGPNCQTNINBCASNPCINOGTCI	807	
Qy	486	THCSSKVDLCLIRPCANGGTCNLNNDYQCTCRAGFTGKDCSDVIDECSSGCPCHNGTCM	545	
Db	808	DDVAGYKNCPLPYTGATCE	827	
Qy	546	NRVNSFEVCANGFRGKQCD	565	
RESULT	10	STANDARD;	PRT;	2531 AA.
ID	NTCL-RAT			
AC	Q07008;			
DT	01-NOV-1995 (REL. 32, CREATED)			
DT	01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)			
DT	01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)			
DE	NEUROGENIC LOCUS NOTCH HOMOLOG PROTEIN 1 PRECURSOR.			
GN	NOTCH1.			
OS	RATTUS NORVEGICUS (RAT).			
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;			
OC	EUTHERIA; RODENTIA.			
RP	SEQUENCE FROM N.A.			
RC	TISSUE-SCHWANN CELL.			
RX	MEDLINE; 92111383.			
RA	WEINMASTER G., ROBERTS V.J., LEMKE G.;			
RL	DEVELOPMENT 113:199-205(1991).			
CC	-!- FUNCTION: REQUIRED FOR THE CORRECT DIFFERENTIATION OF A NUMBER			
CC	OF TISSUES.			
CC	-!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.			
CC	-!- DEVELOPMENTAL STAGE: IN THE EMBRYO, HIGHEST LEVELS OCCUR BETWEEN			
CC	DAYS 12 AND 14 AND DECREASE RAPIDLY TO MUCH LOWER LEVELS IN THE			
CC	ADULT.			
CC	-!- SIMILARITY: HIGH, WITH OTHER NOTCH-TYPE PROTEINS.			
CC	-!- SIMILARITY: CONTAINS 36 EGF-LIKE DOMAINS.			
CC	-!- SIMILARITY: THIS PROTEIN INCLUDES 3 COPIES OF A LIN/NOTCH REPEAT.			
CC	-!- SIMILARITY: CONTAINS 6 ANK REPEATS.			
DR	EMBL; X57405; G57635; .			
DR	PROSITE; PS00010; ASX_HYDROXYL; 22.			
DR	PROSITE; PS00022; EGF_1; 35.			
DR	PROSITE; PS01186; EGF_2; 26.			
DR	PROSITE; PS01187; EGF_CA; 21.			
DR	PROSITE; PS50039; FORK HEAD_3; 1.			
KW	DIFFERENTIATION; NEUROGENESIS; REPEAT; ANK REPEAT; EGF-LIKE DOMAIN;			
KW	TRANSMEMBRANE; SIGNAL; GLYCOPROTEIN.			
FT	SIGNAL 1 18			
FT	CHAIN 19 2531			
FT	DOMAIN 19 1723			
FT	TRANSMEM 1724 1746			
FT	POTENTIAL.			
FT	POTENTIAL.			
FT	NEUROGENIC LOCUS NOTCH HOMOLOG PROTEIN 1.			
FT	EXTRACELLULAR (POTENTIAL).			
FT	POTENTIAL.			

FT	DISULFID	381	398	BY SIMILARITY.
FT	DISULFID	400	409	BY SIMILARITY.
FT	DISULFID	416	429	BY SIMILARITY.
FT	DISULFID	423	438	BY SIMILARITY.
FT	DISULFID	440	449	BY SIMILARITY.
FT	DISULFID	456	467	BY SIMILARITY.
FT	DISULFID	461	476	BY SIMILARITY.
FT	DISULFID	478	487	BY SIMILARITY.
FT	DISULFID	494	505	BY SIMILARITY.
FT	DISULFID	499	514	BY SIMILARITY.
FT	DISULFID	516	525	BY SIMILARITY.
FT	DISULFID	532	543	BY SIMILARITY.
FT	DISULFID	537	552	BY SIMILARITY.
FT	DISULFID	554	563	BY SIMILARITY.
FT	DISULFID	570	580	BY SIMILARITY.
FT	DISULFID	575	589	BY SIMILARITY.
FT	DISULFID	591	600	BY SIMILARITY.
FT	DISULFID	607	618	BY SIMILARITY.
FT	DISULFID	612	627	BY SIMILARITY.
FT	DISULFID	629	638	BY SIMILARITY.
FT	DISULFID	645	655	BY SIMILARITY.
FT	DISULFID	650	664	BY SIMILARITY.
FT	DISULFID	666	675	BY SIMILARITY.
FT	DISULFID	682	693	BY SIMILARITY.
FT	DISULFID	687	702	BY SIMILARITY.
FT	DISULFID	704	713	BY SIMILARITY.
FT	DISULFID	720	730	BY SIMILARITY.
FT	DISULFID	725	739	BY SIMILARITY.
FT	DISULFID	741	750	BY SIMILARITY.
FT	DISULFID	757	768	BY SIMILARITY.
FT	DISULFID	762	777	BY SIMILARITY.
FT	DISULFID	779	788	BY SIMILARITY.
FT	DISULFID	795	806	BY SIMILARITY.
FT	DISULFID	800	815	BY SIMILARITY.
FT	DISULFID	817	826	BY SIMILARITY.
FT	DISULFID	833	844	BY SIMILARITY.
FT	DISULFID	838	855	BY SIMILARITY.
FT	DISULFID	857	866	BY SIMILARITY.
FT	DISULFID	873	884	BY SIMILARITY.
FT	DISULFID	878	893	BY SIMILARITY.
FT	DISULFID	895	904	BY SIMILARITY.
FT	DISULFID	911	922	BY SIMILARITY.
FT	DISULFID	916	931	BY SIMILARITY.
FT	DISULFID	933	942	BY SIMILARITY.
FT	DISULFID	987	998	BY SIMILARITY.
FT	DISULFID	992	1007	BY SIMILARITY.
FT	DISULFID	1009	1018	BY SIMILARITY.
FT	DISULFID	1025	1036	BY SIMILARITY.
FT	DISULFID	1030	1045	BY SIMILARITY.
FT	DISULFID	1047	1056	BY SIMILARITY.
FT	DISULFID	1063	1074	BY SIMILARITY.
FT	DISULFID	1068	1083	BY SIMILARITY.
FT	DISULFID	1085	1094	BY SIMILARITY.
FT	DISULFID	1101	1122	BY SIMILARITY.
FT	DISULFID	1116	1131	BY SIMILARITY.
FT	DISULFID	1133	1142	BY SIMILARITY.
FT	DISULFID	1149	1160	BY SIMILARITY.
FT	DISULFID	1154	1169	BY SIMILARITY.
FT	DISULFID	1171	1180	BY SIMILARITY.
FT	DISULFID	1187	1198	BY SIMILARITY.
FT	DISULFID	1192	1207	BY SIMILARITY.
FT	DISULFID	1209	1218	BY SIMILARITY.
FT	DISULFID	1225	1244	BY SIMILARITY.
FT	DISULFID	1238	1253	BY SIMILARITY.
FT	DISULFID	1255	1264	BY SIMILARITY.
FT	DISULFID	1271	1284	BY SIMILARITY.
FT	DISULFID	1276	1293	BY SIMILARITY.
FT	DISULFID	1295	1304	BY SIMILARITY.
FT	DISULFID	1311	1322	BY SIMILARITY.
FT	DISULFID	1316	1334	BY SIMILARITY.
FT	DISULFID	1336	1345	BY SIMILARITY.
FT	DISULFID	1352	1363	BY SIMILARITY.
FT	DISULFID	1357	1372	BY SIMILARITY.

FT DISULFID 129 BY SIMILARITY.
FT DISULFID 147 BY SIMILARITY.
FT DISULFID 156 BY SIMILARITY.
FT DISULFID 163 BY SIMILARITY.
FT DISULFID 169 BY SIMILARITY.
FT DISULFID 186 BY SIMILARITY.
FT DISULFID 202 BY SIMILARITY.
FT DISULFID 207 BY SIMILARITY.
FT DISULFID 223 BY SIMILARITY.
FT DISULFID 225 BY SIMILARITY.
FT DISULFID 241 BY SIMILARITY.
FT DISULFID 246 BY SIMILARITY.
FT DISULFID 263 BY SIMILARITY.
FT DISULFID 279 BY SIMILARITY.
FT DISULFID 286 BY SIMILARITY.
FT DISULFID 303 BY SIMILARITY.
FT DISULFID 319 BY SIMILARITY.
FT DISULFID 324 BY SIMILARITY.
FT DISULFID 341 BY SIMILARITY.
FT DISULFID 356 BY SIMILARITY.
FT DISULFID 361 BY SIMILARITY.
FT DISULFID 380 BY SIMILARITY.
FT DISULFID 396 BY SIMILARITY.
FT DISULFID 403 BY SIMILARITY.
FT DISULFID 420 BY SIMILARITY.
FT DISULFID 436 BY SIMILARITY.
FT DISULFID 441 BY SIMILARITY.
FT DISULFID 458 BY SIMILARITY.
FT DISULFID 474 BY SIMILARITY.
FT DISULFID 479 BY SIMILARITY.
FT DISULFID 496 BY SIMILARITY.
FT DISULFID 512 BY SIMILARITY.
FT DISULFID 517 BY SIMILARITY.
FT DISULFID 534 BY SIMILARITY.
FT DISULFID 550 BY SIMILARITY.
FT DISULFID 555 BY SIMILARITY.
FT DISULFID 571 BY SIMILARITY.
FT DISULFID 587 BY SIMILARITY.
FT DISULFID 592 BY SIMILARITY.
FT DISULFID 609 BY SIMILARITY.
FT DISULFID 625 BY SIMILARITY.
FT DISULFID 630 BY SIMILARITY.
FT DISULFID 646 BY SIMILARITY.
FT DISULFID 662 BY SIMILARITY.
FT DISULFID 667 BY SIMILARITY.
FT DISULFID 684 BY SIMILARITY.
FT DISULFID 700 BY SIMILARITY.
FT DISULFID 705 BY SIMILARITY.
FT DISULFID 721 BY SIMILARITY.
FT DISULFID 739 BY SIMILARITY.
FT DISULFID 744 BY SIMILARITY.
FT DISULFID 761 BY SIMILARITY.
FT DISULFID 776 BY SIMILARITY.
FT DISULFID 781 BY SIMILARITY.
FT DISULFID 799 BY SIMILARITY.
FT DISULFID 815 BY SIMILARITY.
FT DISULFID 821 BY SIMILARITY.
FT DISULFID 838 BY SIMILARITY.
FT DISULFID 854 BY SIMILARITY.
FT DISULFID 859 BY SIMILARITY.
FT DISULFID 876 BY SIMILARITY.
FT DISULFID 892 BY SIMILARITY.
FT DISULFID 897 BY SIMILARITY.
FT DISULFID 913 BY SIMILARITY.
FT DISULFID 929 BY SIMILARITY.
FT DISULFID 934 BY SIMILARITY.
FT DISULFID 951 BY SIMILARITY.
FT DISULFID 967 BY SIMILARITY.
FT DISULFID 972 BY SIMILARITY.
FT DISULFID 989 BY SIMILARITY.
FT DISULFID 1005 BY SIMILARITY.
FT DISULFID 1010 BY SIMILARITY.
FT DISULFID 1025 BY SIMILARITY.
FT DISULFID 1041 BY SIMILARITY.
FT DISULFID 1056 BY SIMILARITY.

FT DISULFID 1073 BY SIMILARITY.
FT DISULFID 1089 BY SIMILARITY.
FT DISULFID 1094 BY SIMILARITY.
FT DISULFID 1111 BY SIMILARITY.
FT DISULFID 1127 BY SIMILARITY.
FT DISULFID 1132 BY SIMILARITY.
FT DISULFID 1149 BY SIMILARITY.
FT DISULFID 1165 BY SIMILARITY.
FT DISULFID 1177 BY SIMILARITY.
FT DISULFID 1194 BY SIMILARITY.
FT DISULFID 1210 BY SIMILARITY.
FT DISULFID 1215 BY SIMILARITY.
FT DISULFID 1235 BY SIMILARITY.
FT DISULFID 1251 BY SIMILARITY.
FT DISULFID 1256 BY SIMILARITY.
FT DISULFID 1278 BY SIMILARITY.
FT DISULFID 1294 BY SIMILARITY.
FT DISULFID 1299 BY SIMILARITY.
FT DISULFID 1316 BY SIMILARITY.
FT DISULFID 1340 BY SIMILARITY.
FT DISULFID 1345 BY SIMILARITY.
FT DISULFID 1364 BY SIMILARITY.
FT CARBOHYD 1180 POTENTIAL.
FT CARBOHYD 1337 POTENTIAL.
FT CARBOHYD 1439 POTENTIAL.
FT CARBOHYD 1723 POTENTIAL.
FT CARBOHYD 1895 POTENTIAL.
FT CARBOHYD 2278 POTENTIAL.
SQ SEQUENCE 2318 AA; 244245 MW; 879EFAF6 CRC32;

Query Match 14.6%; Score 898; DB 1; Length 2318;
Best Local Similarity 37.7%; Pred. No. 1.75e-189;
Matches 143; Conservative 74; Mismatches 125; Indels 37; Gaps 29;

Db 302 VCVNGWTGSCSQNIDDCATVCFHGATCHDRVASFYACPMGKTGLLCHLDDACVSNP- 360
QY 216 ICLTGMQDYC-H-IPKCAKG-CEHGH-C-DK-PN-QCVCQLGKGCALC--NE-CVLEPN 265
Db 361 CHEDAICDTNPVSGRAICTCPGFTGGACDDVDECSIGANPCEHLGRCVNT-QGSFLCQ 419
QY 266 C-IHGTCN-KP---WT-CICNEGWGLYCQDLNYCT-NHRPCKNGGTCTFNTGEGLYTK 318
Db 420 CGRGYTPRCETDVNEC-LS-GPCRNQATCLD--RI-GOFTCICMAGFTGYCEVDIDE 474
QY 319 CAPGYSGDCENEIYSCDADVNPCQNGGTCTDEPHYTKYKCHANGWGKMCCEKVLTC 378
Db 475 QSSPCVNGGVCKD-R--V--NG--FSCTCPGSGFSGMQLDVDECASTPCRNGAKCVDQP 527
QY 379 SDKPC-HQIGICRNVRLGSKGQGYOCECPIGYSGPNCDLQLDNCSNPNCINGGSC--QP 435
Db 528 DGYECCRCAGFEGTLCERNVDDCSPPCHHG-RCVDGIAFSCACAPGYTGIRCESQVDE 586
QY 436 SG-KCICPAGFSGTRCTNIDDCLGHCQENGGTCTDMVNVQYRCQCPGFGHGHCSKVDL 494
Db 587 CRSQPCRYGKCLDLYKVLRCPCPGTTGVNCEVNIIDDCASNPC-TFGVCRDGINRYDCV 645
QY 495 CLIRPCANGGTCLNLDNDYQCTCRAGTGRGDCSDIDECSGPGCHNGGTCTMNRVNSFEV 554
Db 646 CQPGFTGPLCNVEINECAS 664
QY 555 CANGFRGKQCDSEYSV 573
RESULT 13
ID FBP3_STRPU STANDARD; PRT; 570 AA.
AC P49013.
DT 01-FEB-1996 (REL. 33, CREATED)
DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE FIBROPELLIN C PRECURSOR (EPIDERMAL GROWTH FACTOR-RELATED PROTEIN 3)
DE (EGF III) (FIBROPELLIN III).
GN EGF3.
OS STRONGYLOCENTROTUS PURPURATUS (PURPLE SEA URCHIN).

Query Match	Best Local Similarity	Score	DB 1: Length
Matches 122; Conservative	43.7%;	881; DB 1: 570;	
	61; Mismatches	78; Indels	18; Gaps 14;

[illegible]

RESULT	14	STANDARD	PRT	1964	AA.
ID	NTCA_MOUSE				
AC	P31695; 062389.				
DT	01-JUL-1993 (REL. 26, CREATED)				
DT	01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)				
DT	01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)				
DE	NEUROGENIC LOCUS NOTCH HOMOLOG PROTEIN 4 PRECURSOR (TRANSFORMING				
DE	PROTEIN INT-3).				
GN	NOTCH4 OR INT3 OR INT-3.				
OS	MUS MUSCULUS (MOUSE).				
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;				
OC	EUTHERIA; RODENTIA.				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE; 92194507.				
RA	ROBBINS J., BLONDEL B.J., GALLAHAN D., CALLAHAN R.;				
RL	J. VIROL. 66:2594-2599(1992).				
RN	[2]				
RP	REVISIONS; SEQUENCE FROM N.A.				
RA	CALLAHAN R.;				
RL	SUBMITTED (NOV-1996) TO EMBL/GENBANK/DBJ DATA BANKS.				
RL	[3]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=LUNG, AND TESTIS;				
RX	MEDLINE; 96281668.				
RA	UTTERDAELE H., MAHAZETI G., WU G., YAN Q., SASSOON D., KITAJEWSKI J.;				
RL	DEVELOPMENT 123:2251-2259(1996).				
CC	-1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.				
CC	-1- DISEASE: ACTIVATED INT-3 TRANSFORMS MAMMARY EPITHELIAL CELLS.				
CC	-1- SIMILARITY: CONTAINS 29 EGF-LIKE DOMAINS.				
CC	-1- SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.				
CC	-1- SIMILARITY: CONTAINS 6 CDC10/SW16 REPEATS.				
CC	-1- SIMILARITY: CONTAINS 6 ANK REPEATS.				
DR	EMBL; M80456; G1714084; -				
DR	EMBL; U43691; G1401160; -				
DR	PIR; A38072; TVMW73.				
DR	MGI; MGI:107471; NOTCH4.				
DR	PROSITE; PS00010; ASX_HYDROXYL, 11.				
DR	PROSITE; PS00022; EGF_1; 28.				
DR	PROSITE; PS01186; EGF_2; 21.				
DR	PROSITE; PS01187; EGF_CA; 9.				
KW	DIFFERENTIATION; NEUROGENESIS; REPEAT; EGF-LIKE DOMAIN; TRANSMEMBRANE				
KW	GLYCOPROTEIN; ONCOGENE; ANK REPEAT; SIGNAL.				
FT	SIGNAL	1	20	POTENTIAL.	
FT	CHAIN	21	1964	NEUROGENIC LOCUS NOTCH HOMOLOG PROTEIN 4	
FT	DOMAIN	21	1443	EXTRACELLULAR (POTENTIAL).	
FT	TRANSMEM	1444	1464	POTENTIAL.	
FT	DOMAIN	1465	1964	CYTOPLASMIC (POTENTIAL).	
FT	DOMAIN	21	60	EGF-LIKE 1.	
FT	DOMAIN	61	112	EGF-LIKE 2.	
FT	DOMAIN	115	152	EGF-LIKE 3.	
FT	DOMAIN	153	189	EGF-LIKE 4.	

```
FT DOMAIN 191 229 EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL). 481 496 BY SIMILARITY.
FT DOMAIN 231 271 EGF-LIKE 6. 498 507 BY SIMILARITY.
FT DOMAIN 309 350 EGF-LIKE 7, CALCIUM-BINDING (POTENTIAL). 514 525 BY SIMILARITY.
FT DOMAIN 311 350 EGF-LIKE 8, CALCIUM-BINDING (POTENTIAL). 534 545 BY SIMILARITY.
FT DOMAIN 352 388 EGF-LIKE 9, CALCIUM-BINDING (POTENTIAL). 536 545 BY SIMILARITY.
FT DOMAIN 389 427 EGF-LIKE 10. 552 563 BY SIMILARITY.
FT DOMAIN 429 470 EGF-LIKE 11, CALCIUM-BINDING (POTENTIAL). 557 572 BY SIMILARITY.
FT DOMAIN 472 508 EGF-LIKE 12, CALCIUM-BINDING (POTENTIAL). 574 583 BY SIMILARITY.
FT DOMAIN 510 546 EGF-LIKE 13, CALCIUM-BINDING (POTENTIAL). 590 601 BY SIMILARITY.
FT DOMAIN 548 584 EGF-LIKE 14, CALCIUM-BINDING (POTENTIAL). 595 610 BY SIMILARITY.
FT DOMAIN 586 622 EGF-LIKE 15, CALCIUM-BINDING (POTENTIAL). 612 621 BY SIMILARITY.
FT DOMAIN 622 656 EGF-LIKE 16. 626 637 BY SIMILARITY.
FT DOMAIN 658 686 EGF-LIKE 17. 631 646 BY SIMILARITY.
FT DOMAIN 688 724 EGF-LIKE 18. 648 655 BY SIMILARITY.
FT DOMAIN 726 762 EGF-LIKE 19. 662 669 BY SIMILARITY.
FT DOMAIN 764 800 EGF-LIKE 20. 664 674 BY SIMILARITY.
FT DOMAIN 803 839 EGF-LIKE 21. 676 685 BY SIMILARITY.
FT DOMAIN 841 877 EGF-LIKE 22. 692 703 BY SIMILARITY.
FT DOMAIN 878 924 EGF-LIKE 23. 697 712 BY SIMILARITY.
FT DOMAIN 926 962 EGF-LIKE 24. 714 723 BY SIMILARITY.
FT DOMAIN 964 1000 EGF-LIKE 25. 730 741 BY SIMILARITY.
FT DOMAIN 1002 1040 EGF-LIKE 26. 735 750 BY SIMILARITY.
FT DOMAIN 1042 1081 EGF-LIKE 27. 752 761 BY SIMILARITY.
FT DOMAIN 1083 1122 EGF-LIKE 28. 768 779 BY SIMILARITY.
FT DOMAIN 1126 1167 EGF-LIKE 29. 773 788 BY SIMILARITY.
FT DOMAIN 1168 1282 3 X LIN/NOTCH REPEATS. 790 799 BY SIMILARITY.
FT REPEAT 1208 1248 LIN/NOTCH 1. 807 818 BY SIMILARITY.
FT REPEAT 1209 1242 LIN/NOTCH 2. 827 838 BY SIMILARITY.
FT REPEAT 1243 1282 LIN/NOTCH 3. 829 838 BY SIMILARITY.
FT DOMAIN 1572 1785 6 X ANK MOTIF REPEATS. 845 856 BY SIMILARITY.
FT REPEAT 1572 1603 ANK MOTIF 1. 850 865 BY SIMILARITY.
FT REPEAT 1622 1653 ANK MOTIF 2. 867 876 BY SIMILARITY.
FT REPEAT 1654 1685 ANK MOTIF 3. 882 903 BY SIMILARITY.
FT REPEAT 1688 1719 ANK MOTIF 4. 897 912 BY SIMILARITY.
FT REPEAT 1721 1752 ANK MOTIF 5. 914 923 BY SIMILARITY.
FT REPEAT 1754 1785 ANK MOTIF 6. 941 941 BY SIMILARITY.
FT DISULFID 25 38 BY SIMILARITY. 935 950 BY SIMILARITY.
FT DISULFID 32 48 BY SIMILARITY. 952 961 BY SIMILARITY.
FT DISULFID 50 59 BY SIMILARITY. 968 979 BY SIMILARITY.
FT DISULFID 65 77 BY SIMILARITY. 973 988 BY SIMILARITY.
FT DISULFID 71 100 BY SIMILARITY. 990 999 BY SIMILARITY.
FT DISULFID 102 111 BY SIMILARITY. 1006 1019 BY SIMILARITY.
FT DISULFID 119 130 BY SIMILARITY. 1011 1028 BY SIMILARITY.
FT DISULFID 124 140 BY SIMILARITY. 1030 1039 BY SIMILARITY.
FT DISULFID 142 151 BY SIMILARITY. 1046 1057 BY SIMILARITY.
FT DISULFID 157 168 BY SIMILARITY. 1051 1069 BY SIMILARITY.
FT DISULFID 162 177 BY SIMILARITY. 1071 1080 BY SIMILARITY.
FT DISULFID 179 188 BY SIMILARITY. 1087 1098 BY SIMILARITY.
FT DISULFID 195 208 BY SIMILARITY. 1092 1110 BY SIMILARITY.
FT DISULFID 202 217 BY SIMILARITY. 1112 1121 BY SIMILARITY.
FT DISULFID 219 228 BY SIMILARITY. 1130 1142 BY SIMILARITY.
FT DISULFID 235 246 BY SIMILARITY. 1136 1155 BY SIMILARITY.
FT DISULFID 240 259 BY SIMILARITY. 1157 1166 BY SIMILARITY.
FT DISULFID 261 270 BY SIMILARITY. 1171 711 POTENTIAL.
FT DISULFID 235 246 BY SIMILARITY. 960 960 POTENTIAL.
FT DISULFID 240 259 BY SIMILARITY. 1139 1139 POTENTIAL.
FT DISULFID 261 270 BY SIMILARITY. 1547 1547 POTENTIAL.
FT DISULFID 277 288 BY SIMILARITY. 43 43 Q -> R (IN REF. 3).
FT DISULFID 282 297 BY SIMILARITY. 298 298 L -> P (IN REF. 3).
FT DISULFID 299 308 BY SIMILARITY. 884 884 M -> K (IN REF. 3).
FT DISULFID 315 329 BY SIMILARITY. 927 927 M -> V (IN REF. 3).
FT DISULFID 323 338 BY SIMILARITY. 966 966 E -> D (IN REF. 3).
FT DISULFID 340 349 BY SIMILARITY. 1245 1245 L -> P (IN REF. 3).
FT DISULFID 356 367 BY SIMILARITY. 1489 1489 A -> T (IN REF. 3).
FT DISULFID 371 376 BY SIMILARITY. 1549 1549 S -> G (IN REF. 3).
FT DISULFID 378 387 BY SIMILARITY. 1688 1688 S -> T (IN REF. 3).
FT DISULFID 393 404 BY SIMILARITY. 1836 1838 AAA -> APR (IN REF. 3).
FT DISULFID 398 415 BY SIMILARITY. 1964 AA; 206691 MW; 3F16B6B5 CRC32;
FT DISULFID 417 426 BY SIMILARITY.
FT DISULFID 433 449 BY SIMILARITY.
FT DISULFID 443 458 BY SIMILARITY.
FT DISULFID 460 469 BY SIMILARITY.
FT DISULFID 476 487 BY SIMILARITY.

Query Match 12.9%; Score 798; DB 1; Length 1964;
Best Local Similarity 37.2%; Pred. No. 7.43e-164;
Matches 123; Conservative 73; Mismatches 110; Indels 25; Gaps 20;
```

Db	270	CEMNPDCCVHRCQNGATCTLDGL-DT-YTCLPRTWKGMDCSEDIDECANGPPRCRNG	327	FT	DOMAIN	388	425	EGF-LIKE 4, CALCIUM-BINDING (POTENTIAL).
QY	248	CQLGKRALCNCEVLEPNCICGTCNKPTCICNGWGLYNODLNTNH-RP-CRKG	305	FT	DOMAIN	427	463	EGF-LIKE 5.
Db	328	TCQNTA-GSFHCVCVSGAGACEENIDCAAT--CAPGSTCID--RV-GSFCLCPG	381	FT	DOMAIN	464	500	EGF-LIKE 6.
QY	306	TCFVTGGLYTKCAKCAPYSGDCENEYISCADVNPCCNGSTCIDEPHTKGYKCHCANG	365	FT	DOMAIN	501	532	EGF-LIKE 7.
Db	382	RTGLLCLLEDM-CLSQCHVNAOC-STNP-L-T-GSTL-CLCQGYSGSTGHODIDECQ	435	FT	DOMAIN	532	581	EGF-LIKE 8.
QY	366	WSGMECEKVLTCSDKPC-H-OGICRNVRPGLSKGQGYCCPCPIGYSGPCDQLQDNC--	422	FT	DOMAIN	581	611	EGF-LIKE 9.
Db	436	AQGPSCCEHGGSCINTPGSFNCLCPGYTSRCEADHNECLSQPCHGSTCLDLATFH	495	FT	DOMAIN	609	646	EGF-LIKE 10.
QY	423	S---PNCINGSC-Q-PSG-KCICPAGFSGTRFETNIDDLGHCENGSTCIDMVNQR	476	FT	DOMAIN	646	685	EGF-LIKE 11, CALCIUM-BINDING (POTENTIAL).
Db	496	CLCPGLEGRICEVEVECTSNPCLNQAACHDLNGFCCLCPGTGARCENKDECSST	555	FT	DOMAIN	685	723	EGF-LIKE 12, CALCIUM-BINDING (POTENTIAL).
QY	477	CQCPVGFHGTSCSKVLDCLIRPCANGTCLINNDYQCTCRAGFTGKDCSVIDECSSG	536	FT	DOMAIN	723	761	EGF-LIKE 13, CALCIUM-BINDING (POTENTIAL).
Db	556	PCANGRCRCROPGAFYCECLPFGFPGHCEKE	586	FT	DOMAIN	761	800	EGF-LIKE 14, CALCIUM-BINDING (POTENTIAL).
QY	537	PCHNGTGMNRVNSFECVCANGFRGKQDDEE	567	FT	DOMAIN	800	838	EGF-LIKE 15, CALCIUM-BINDING (POTENTIAL).
RESULT	15	STANDARD:	PRT: 2139 AA.					
ID	CRB.DROME			FT	DOMAIN	838	890	EGF-LIKE 16, CALCIUM-BINDING (POTENTIAL).
AC	P10040;			FT	DOMAIN	890	929	EGF-LIKE 17, CALCIUM-BINDING (POTENTIAL).
DT	01-MAR-1989 (REL. 10, CREATED)			FT	DOMAIN	929	978	EGF-LIKE 18, CALCIUM-BINDING (POTENTIAL).
DT	01-MAY-1991 (REL. 18, LAST SEQUENCE UPDATE)			FT	DOMAIN	978	1021	EGF-LIKE 19.
DT	01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)			FT	DOMAIN	1021	1243	EGF-LIKE 20.
DE	CRUMBS PROTEIN PRECURSOR (95F).			FT	DOMAIN	1243	1517	EGF-LIKE 21.
GN	CRB.			FT	DOMAIN	1517	1795	EGF-LIKE 22.
OS	DROSOPHILA MELANOGASTER (FRUIT FLY).			FT	DOMAIN	1795	1833	EGF-LIKE 23.
OC	EUKARYOTA; METAZOA; ARTHROPODA; INSECTA; DIPTERA.			FT	DOMAIN	1833	1871	EGF-LIKE 24, CALCIUM-BINDING (POTENTIAL).
OC	[1]			FT	DOMAIN	1871	1915	EGF-LIKE 25.
RC	SEQUENCE FROM N.A.			FT	DOMAIN	1915	1951	EGF-LIKE 26.
RX	STRAIN-OREGON-R; TISSUE-EMBRYO;			FT	DOMAIN	1951	1991	EGF-LIKE 27, CALCIUM-BINDING (POTENTIAL).
RA	TEPASS U., THERES C., KNUST E.;			FT	DOMAIN	1991	2029	EGF-LIKE 28, CALCIUM-BINDING (POTENTIAL).
RL	CELL 61:787-799(1990).			FT	DOMAIN	2029	2070	EGF-LIKE 29.
RL	[2]			FT	DOMAIN	2070	282	BY SIMILARITY.
RN	SEQUENCE OF 1663-1955 FROM N.A.			FT	DOMAIN	282	291	BY SIMILARITY.
RX	MEDLINE; 87218537.			FT	DOMAIN	291	302	BY SIMILARITY.
RA	KNUST E., DIETRICH U., TEPASS U., BREMER K.A., WEIGEL D.,			FT	DOMAIN	302	321	BY SIMILARITY.
RA	VAESSIN H., CAMPOS-ORTEGA J.A.;			FT	DOMAIN	321	331	BY SIMILARITY.
RL	EMBO J. 6:761-766(1987).			FT	DOMAIN	331	342	BY SIMILARITY.
RN	[2]			FT	DOMAIN	342	363	BY SIMILARITY.
RP	SEQUENCE OF 1663-1955 FROM N.A.			FT	DOMAIN	363	374	BY SIMILARITY.
RA	KNUST E., DIETRICH U., TEPASS U., BREMER K.A., WEIGEL D.,			FT	DOMAIN	374	385	BY SIMILARITY.
RA	VAESSIN H., CAMPOS-ORTEGA J.A.;			FT	DOMAIN	385	403	BY SIMILARITY.
RL	EMBO J. 6:761-766(1987).			FT	DOMAIN	403	412	BY SIMILARITY.
RN	[2]			FT	DOMAIN	412	424	BY SIMILARITY.
RP	SEQUENCE OF 1663-1955 FROM N.A.			FT	DOMAIN	424	442	BY SIMILARITY.
RA	KNUST E., DIETRICH U., TEPASS U., BREMER K.A., WEIGEL D.,			FT	DOMAIN	442	451	BY SIMILARITY.
RA	VAESSIN H., CAMPOS-ORTEGA J.A.;			FT	DOMAIN	451	462	BY SIMILARITY.
RL	EMBO J. 6:761-766(1987).			FT	DOMAIN	462	479	BY SIMILARITY.
RN	[2]			FT	DOMAIN	479	488	BY SIMILARITY.
RP	SEQUENCE OF 1663-1955 FROM N.A.			FT	DOMAIN	488	499	BY SIMILARITY.
RA	KNUST E., DIETRICH U., TEPASS U., BREMER K.A., WEIGEL D.,			FT	DOMAIN	499	515	BY SIMILARITY.
RA	VAESSIN H., CAMPOS-ORTEGA J.A.;			FT	DOMAIN	515	520	BY SIMILARITY.
RL	EMBO J. 6:761-766(1987).			FT	DOMAIN	520	531	BY SIMILARITY.
RN	[2]			FT	DOMAIN	531	562	BY SIMILARITY.
RP	SEQUENCE OF 1663-1955 FROM N.A.			FT	DOMAIN	562	569	BY SIMILARITY.
RA	KNUST E., DIETRICH U., TEPASS U., BREMER K.A., WEIGEL D.,			FT	DOMAIN	569	580	BY SIMILARITY.
RA	VAESSIN H., CAMPOS-ORTEGA J.A.;			FT	DOMAIN	580	597	BY SIMILARITY.
RL	EMBO J. 6:761-766(1987).			FT	DOMAIN	597	602	BY SIMILARITY.
RN	[2]			FT	DOMAIN	602	610	BY SIMILARITY.
RP	SEQUENCE OF 1663-1955 FROM N.A.			FT	DOMAIN	610	624	BY SIMILARITY.
RA	KNUST E., DIETRICH U., TEPASS U., BREMER K.A., WEIGEL D.,			FT	DOMAIN	624	634	BY SIMILARITY.
RA	VAESSIN H., CAMPOS-ORTEGA J.A.;			FT	DOMAIN	634	645	BY SIMILARITY.
RL	EMBO J. 6:761-766(1987).			FT	DOMAIN	645	664	BY SIMILARITY.
RN	[2]			FT	DOMAIN	664	673	BY SIMILARITY.
RP	SEQUENCE OF 1663-1955 FROM N.A.			FT	DOMAIN	673	684	BY SIMILARITY.
RA	KNUST E., DIETRICH U., TEPASS U., BREMER K.A., WEIGEL D.,			FT	DOMAIN	684	702	BY SIMILARITY.
RA	VAESSIN H., CAMPOS-ORTEGA J.A.;			FT	DOMAIN	702	711	BY SIMILARITY.
RL	EMBO J. 6:761-766(1987).			FT	DOMAIN	711	722	BY SIMILARITY.
RN	[2]			FT	DOMAIN	722	740	BY SIMILARITY.
RP	SEQUENCE OF 1663-1955 FROM N.A.			FT	DOMAIN	740	749	BY SIMILARITY.
RA	KNUST E., DIETRICH U., TEPASS U., BREMER K.A., WEIGEL D.,			FT	DOMAIN	749	760	BY SIMILARITY.
RA	VAESSIN H., CAMPOS-ORTEGA J.A.;			FT	DOMAIN	760	778	BY SIMILARITY.
RL	EMBO J. 6:761-766(1987).			FT	DOMAIN	778	787	BY SIMILARITY.
RN	[2]			FT	DOMAIN	787	799	BY SIMILARITY.
RP	SEQUENCE OF 1663-1955 FROM N.A.			FT	DOMAIN	799	817	BY SIMILARITY.
RA	KNUST E., DIETRICH U., TEPASS U., BREMER K.A., WEIGEL D.,			FT	DOMAIN	817	826	BY SIMILARITY.
RA	VAESSIN H., CAMPOS-ORTEGA J.A.;			FT	DOMAIN	826	837	BY SIMILARITY.
RL	EMBO J. 6:761-766(1987).			FT	DOMAIN	837	855	BY SIMILARITY.
RN	[2]			FT	DOMAIN	855	890	BY SIMILARITY.
RP	SEQUENCE OF 1663-1955 FROM N.A.			FT	DOMAIN	890		BY SIMILARITY.

DISULFID 892 BY SIMILARITY.
DISULFID 908 BY SIMILARITY.
DISULFID 913 BY SIMILARITY.
DISULFID 928 BY SIMILARITY.
DISULFID 939 BY SIMILARITY.
DISULFID 946 BY SIMILARITY.
DISULFID 957 BY SIMILARITY.
DISULFID 966 BY SIMILARITY.
DISULFID 968 BY SIMILARITY.
DISULFID 977 BY SIMILARITY.
DISULFID 984 BY SIMILARITY.
DISULFID 989 BY SIMILARITY.
DISULFID 1009 BY SIMILARITY.
DISULFID 1020 BY SIMILARITY.
DISULFID 1021 BY SIMILARITY.
DISULFID 1211 BY SIMILARITY.
DISULFID 1216 BY SIMILARITY.
DISULFID 1231 BY SIMILARITY.
DISULFID 1233 BY SIMILARITY.
DISULFID 1242 BY SIMILARITY.
DISULFID 1496 BY SIMILARITY.
DISULFID 1498 BY SIMILARITY.
DISULFID 1505 BY SIMILARITY.
DISULFID 1507 BY SIMILARITY.
DISULFID 1516 BY SIMILARITY.
DISULFID 1763 BY SIMILARITY.
DISULFID 1768 BY SIMILARITY.
DISULFID 1783 BY SIMILARITY.
DISULFID 1785 BY SIMILARITY.
DISULFID 1801 BY SIMILARITY.
DISULFID 1806 BY SIMILARITY.
DISULFID 1821 BY SIMILARITY.
DISULFID 1832 BY SIMILARITY.
DISULFID 1839 BY SIMILARITY.
DISULFID 1850 BY SIMILARITY.
DISULFID 1844 BY SIMILARITY.
DISULFID 1861 BY SIMILARITY.
DISULFID 1870 BY SIMILARITY.
DISULFID 1878 BY SIMILARITY.
DISULFID 1883 BY SIMILARITY.
DISULFID 1905 BY SIMILARITY.
DISULFID 1919 BY SIMILARITY.
DISULFID 1924 BY SIMILARITY.
DISULFID 1939 BY SIMILARITY.
DISULFID 1941 BY SIMILARITY.
DISULFID 1957 BY SIMILARITY.
DISULFID 1968 BY SIMILARITY.
DISULFID 1979 BY SIMILARITY.
DISULFID 1988 BY SIMILARITY.
DISULFID 1995 BY SIMILARITY.
DISULFID 2002 BY SIMILARITY.
DISULFID 2017 BY SIMILARITY.
DISULFID 2019 BY SIMILARITY.
CARBOHYD 37 POTENTIAL.
CARBOHYD 96 POTENTIAL.
CARBOHYD 198 POTENTIAL.
CARBOHYD 238 POTENTIAL.
CARBOHYD 239 POTENTIAL.
CARBOHYD 239 POTENTIAL.
CARBOHYD 336 POTENTIAL.
CARBOHYD 400 POTENTIAL.
CARBOHYD 550 POTENTIAL.
CARBOHYD 565 POTENTIAL.
CARBOHYD 736 POTENTIAL.
CARBOHYD 746 POTENTIAL.
CARBOHYD 860 POTENTIAL.
CARBOHYD 884 POTENTIAL.
CARBOHYD 976 POTENTIAL.
CARBOHYD 1102 POTENTIAL.
CARBOHYD 1114 POTENTIAL.
CARBOHYD 1138 POTENTIAL.
CARBOHYD 1192 POTENTIAL.
CARBOHYD 1245 POTENTIAL.
CARBOHYD 1255 POTENTIAL.
CARBOHYD 1354 POTENTIAL.
CARBOHYD 1363 POTENTIAL.
CARBOHYD 1441 POTENTIAL.
CARBOHYD 1454 POTENTIAL.
CARBOHYD 1546 POTENTIAL.
CARBOHYD 1740 POTENTIAL.
CARBOHYD 1752 POTENTIAL.
CARBOHYD 1809 POTENTIAL.
CARBOHYD 1849 POTENTIAL.
CARBOHYD 1885 POTENTIAL.
CARBOHYD 1894 POTENTIAL.
CARBOHYD 1900 POTENTIAL.
CARBOHYD 2029 POTENTIAL.
CARBOHYD 2035 POTENTIAL.
CARBOHYD 2068 POTENTIAL.
CONFLICT 1663

MYDTPAFSTDIDQ -> WWIRQLFPRTSK (IN

FT CONFLICT 1722 1722 REF. 2).
FT CONFLICT 1955 1955 L -> V (IN REF. 2).
SQ SEQUENCE 2139 AA; 233619 MW; 2B9D/CB6 CRC32;
Query Match 12.4%; Score 765; DB 1; Length 2139;
Best Local Similarity 30.8%; Pred. No. 1.93e-155;
Matches 140; Conservative 96; Mismatches 184; Indels 35; Gaps 29;
Db 450 TCQCLDNGWGETCDRPMTCQTQQ-CFNGGTCLDKPFGFQCLCPPEVTGELCQTAPSCAQ 508
QY 183 TCD-LNYGSG-CAK-F-CRPRDDSFHSTCSET--G-EIICLTGQGDYCHI-PRCAKG 234
Db 509 CPIDSECV-GGKVCVKPGSSGYNCQSTGTGASALAL--TPINCNATNGKCLNGTCSMN 565
QY 235 CE-HGHCDPQNCVQCLGWKALCNECVLEPNCIHGTCKNKPWTCICNEG--WGGLYCNQD 291
Db 566 GTHCYCAVGY-SGDRCEKAENCSPLNCOBPMVCVQNCCLCPENKVCNOCATOPCONGGEC 624
QY 292 LNYCTNHRPCKNGGTCFNTGE-GLYTCKCAPGYSGDDCE-NEIYSCDA-DVNPQNGGTC 348
Db 625 VDLPNG--DYECKCTRGWTGTCGNDVDECT--L-HPKTCGNIGICK-NEKG-SYKCYCTP 677
QY 349 IDEPHRTKTYKCHCANGSGKCEKVLTCSDRCPCHQGICRNVPRGLSGKQGYQCECPI 408
Db 678 GFTGVHCDSDVDECLSPCLNGATCHNKINAYECVCPQGYEGENCEVDIDECGNSPCSNG 737
QY 409 GYSGPNCDLQDNCSNPNCINGSCQP--SG-KCICPAGFSGTRCTEINDCGLHOCENG 465
Db 738 STCIDRINNFTNCIPGMRGRICDIDIDCVGDPCLNGGQCIDQLGGFRDCSGTGYEGE 797
QY 466 GTCIDMVNRYRCQVPGFHTGTHCSSKVDLCLIRPCANGGTCLNNDYQCTCR-AGFTGK 524
Db 798 NCELNIDECLSNPCTNGAKCLDRVKDYFDCDHNGYKGNKEQDINECES-NPCQYNGNCL 856
QY 525 DCSVDIDECSGPGCHNGGTCMNRVNSFEVCANGFRGKQCDSESYDVTFDAHQYGATTQ 584
Db 857 ERSNITLYQMSRITDLPKVFSQPFSENASGYECV 891
QY 585 ARAD-GLTN-AQVV-LIAVFSVAMPLVAVIA-ACV 615
Search completed: Fri Jun 12 11:55:14 1998
Job time : 92 secs.

FT SIGNAL 1 25 POTENTIAL.
FT CHAIN 26 832 DELTA PROTEIN.
FT DOMAIN 26 595 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 569 617 POTENTIAL.
FT DOMAIN 618 832 INTRACELLULAR (POTENTIAL).
FT DOMAIN 217 256 9 EGF-TYPE REPEATS.
FT REPEAT 217 256 EGF-LIKE 1.
FT REPEAT 257 290 EGF-LIKE 2.
FT REPEAT 291 330 EGF-LIKE 3.
FT REPEAT 331 373 EGF-LIKE 4.
FT REPEAT 374 417 EGF-LIKE 5.
FT REPEAT 418 452 EGF-LIKE 6.
FT REPEAT 453 490 EGF-LIKE 7.
FT REPEAT 491 528 EGF-LIKE 8.
FT REPEAT 529 566 EGF-LIKE 9.
FT CARBOHYD 98 98 POTENTIAL.
FT CARBOHYD 137 137 POTENTIAL.
FT CARBOHYD 167 167 POTENTIAL.
FT CARBOHYD 421 421 POTENTIAL.
FT CARBOHYD 649 649 POTENTIAL.
FT CONFLICT 437 438 GK -> ET (IN REF. 2).
FT CONFLICT 443 443 A -> S (IN REF. 2).
FT CONFLICT 459 459 G -> A (IN REF. 2).
FT CONFLICT 490 490 S -> T (IN REF. 2).
FT CONFLICT 591 591 T -> A (IN REF. 2).
SQ SEQUENCE 832 AA; 88813 MW; CF9ABEC1 CRC32;

Query Match 99.6%; Score 6140; DB 3; Length 832;
Best Local Similarity 99.6%; Pred. No. 0.00e+00;
Matches 830; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

Db 1 MHWIKCLTAFCTFVIVQVHSSGSEFELRLKYFSDHGRDNEGRCCSGSDGATGKCLGS 60
QY 1 MHWIKCLTAFCTFVIVQVHSSGSEFELRLKYFSDHGRDNEGRCCSGSDGATGKCLGS 60
Db 61 KTRFRVCLKHQOATIDTTSQCTYGDVITPILGENSVNLTDQORFONKFTNPQIPFPFF 120
QY 61 KTRFRVCLKHQOATIDTTSQCTYGDVITPILGENSVNLTDQORFONKFTNPQIPFPFF 120
Db 121 SWPGTFSLIVEAWHTNNSGNARTNKLRLQVQVLEVSSEWTKNKSQYTSLEYDF 180
QY 121 SWPGTFSLIVEAWHTNNSGNARTNKLRLQVQVLEVSSEWTKNKSQYTSLEYDF 180
Db 181 RVTCDLNTYSGCAFCRPRDPSFGHSTCSETGEIICLTGWQGDYCHIPKACGCEHGH 240
QY 181 RVTCDLNTYSGCAFCRPRDPSFGHSTCSETGEIICLTGWQGDYCHIPKACGCEHGH 240
Db 241 DKPNOCVQLGKALCNCEVLEPCNPIHGTCKNKPWTCICNEGWGLYCNODLNYCTNHRP 300
QY 241 DKPNOCVQLGKALCNCEVLEPCNPIHGTCKNKPWTCICNEGWGLYCNODLNYCTNHRP 300
Db 301 CKNGTCTNFBGLYTKCAPYSGDDCENETSCDADVPNCQNGGTCIDEPHTKTGYKC 360
QY 301 CKNGTCTNFBGLYTKCAPYSGDDCENETSCDADVPNCQNGGTCIDEPHTKTGYKC 360
Db 361 HCANGWSKMCCEKVLTSCKPCHOGICRNPVPLGSKGQYQCECPYIGSGPNCDLQLD 420
QY 361 HCANGWSKMCCEKVLTSCKPCHOGICRNPVPLGSKGQYQCECPYIGSGPNCDLQLD 420
Db 421 NCSPNPLNGGSCQPSGKICIPAGFSGTCTNIDCLHQCENGTCIDMNVQYRCQCV 480
QY 421 NCSPNPLNGGSCQPSGKICIPAGFSGTCTNIDCLHQCENGTCIDMNVQYRCQCV 480
Db 481 PGFHTGTHSSKVDLCLIRPCANGGTCLNLDNDYQCTCRAGFTGKDCSVDIDESGPGCHN 540
QY 481 PGFHTGTHSSKVDLCLIRPCANGGTCLNLDNDYQCTCRAGFTGKDCSVDIDESGPGCHN 540
Db 541 GGTCHNRVNSFCVCANGFRGQCDEESYDSVTFDAHQYGATTQARADGLTNAQVYLI 600
QY 541 GGTCHNRVNSFCVCANGFRGQCDEESYDSVTFDAHQYGATTQARADGLTNAQVYLI 600
Db 601 FSVAMPLVAVIAACVFCMKRKRKRAQEKDDAEARKQNEQNAVATMHNGSGVGVVALASA 660

QY 601 FSVAMPLVAVIAACVFCMKRKRKRAQEKDDAEARKQNEQNAVATMHNGSGVGVVALASA 660
Db 661 SLGGKTSNSGLTFDGGNPNIIKNTWDSVNNICASAAAAAADADECLMTGGVVASV 720
QY 661 SLGGKTSNSGLTFDGGNPNIIKNTWDSVNNICASAAAAAADADECLMTGGVVASV 720
Db 721 ADNNNANSDFCVAPLQRAKSOQLNTDPTLMHRGSPAGSSAKGASGGGPGGAAGKRISVL 780
QY 721 ADNNNANSDFCVAPLQRAKSOQLNTDPTLMHRGSPAGSSAKGASGGGPGGAAGKRISVL 780
Db 781 GEGSYCSQRPWSLAAAGVAGACSSQLMAAASAAAGTD-GTAQOQSRVVCCTGPHM 832
QY 781 GEGSYCSQRPWSLAAAGVAGACSSQLMAAASAAAGTAGTAQOQSRVVCCTGPHM 833

RESULT 2
ID Q90656 PRELIMINARY; PRT; 728 AA.
AC Q90656;
DT 01-NOV-1996 (TREMBREL. 01, CREATED)
DT 01-NOV-1996 (TREMBREL. 01, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBREL. 05, LAST ANNOTATION UPDATE)
DE TRANSMEMBRANE PROTEIN C-DELTA-1.
OS GALLUS GALLUS (CHICKEN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AVES; NEOGNATHAE;
OC GALLIFORMES.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-SPINAL CORD.
RA MEDLINE; 95319507.
RX HENRIQUE D., ADAM J., MYAT A., CHITNIS A., LEWIS J., ISH-HOROWICZ D.;
RL NATURE 375:787-790(1995).
DR EMBL; U26590; G882412;
DR PROSITE; PS01187; EGF_CA; 2.
KW GLYCOPROTEIN; EGF-LIKE DOMAIN.
SQ SEQUENCE 728 AA; 79861 MW; 7439F575 CRC32;

Query Match 31.0%; Score 1910; DB 12; Length 728;
Best Local Similarity 50.1%; Pred. No. 0.00e+00;
Matches 263; Conservative 78; Mismatches 168; Indels 16; Gaps 15;

Db 20 QVDSGSGVFLKLOEFYNNKGLLSNRNCCGGPGGAGQQCCDKTFFRVCLKHQASVSP 79
QY 19 QVHSSGSEFELRLKYFSDHGRDNEGRCCSGSDGATGKCLGSKCTRFVCLKHQATIDT 78
Db 80 EPCTYGSATPVLGANSFVDPGAGADAFNSPIRPFPGTWPFTSLIIEALHTDSP 139
QY 79 TSQCTYGDVITPILGENSVNLTDQORFONKFTNPQIPFPFSWPGTFSLIVEAWH-DTN 137
Db 140 DDLTTPENPERLISRLATORHLAVGEEMSDLHSSGRDGLKYSYRFVCDHYHGGGVFC 199
QY 138 NSGNARTNKLRLQVQVLEVSSEWTKNKSQYTSLEYDFRVTCDLNTYSGCAKFC 197
Db 200 RPDRDRFGHTGGERGEKVCNPGWKQYCTEPICLPGCDQHQGFCDKPGECKCRVWQGR 259
QY 198 RPDRDPSFGHSTCSETGEIICLTGWQGDYCHIPKACGC-E-HGCDKPNOCVQLGKGA 255
Db 260 YDCETIRYPCGLHGTCCQPMQCNQCBEGWGLFCNDLNYCTHHKPKNGKATCTNTGGSY 319
QY 256 LCNECVLEPCNPIHGTCKNKPWTCICNEGWGLYCNODLNYCTNHRPCKNGTCTNTEGGLY 315
Db 320 TCSRCRGYTGSSCEIINECDA--NPCKNGSGCTD-LE-NS-YSCCTCPPGFYCKNCLSA 374
QY 316 TCRCAPYSGDDCENETSCDADVPNCQNGGTCIDEPHTKTGYKCHCANWSKMCCEKV 375
Db 375 MTCADGPGCFNGG-R-C-TD-NPDC-GYSCRCPLGYSGFNCEKKIDYCSSSPCANGAQCV 429
QY 376 LTCSDAPCHOGICRNPVPLGSKGQYQCECPYIGSGPNCDLQDNCSPNPLNGGSC-Q 434
Db 430 LGNSYICQOAGFTGRHCDNDNDVDCAFPCVNGGTGQDGVNDYSCTCPCPGYKNGKSTPV 489
QY 435 -PSG-KCICPAGFSGTCTNIDCLHQCENGTCIDMNVQYRCQCVPGFHTGTHSSKV 492
Db 490 SRCEHPNCHGATCHERSNRYVCECARGYGLNQCQLLLPEPPQGP 534

Mon Jun 15 11:54:33 1998

493 DLGIRPCANGTCLNNDYQCTCRAGFTGKDCSVYDIDECSSGP 537

RESULT 3 PRELIMINARY; PRT; 717 AA.

ID P87357; PRELIMINARY; PRT; 717 AA.

AC P87357; PRELIMINARY; PRT; 717 AA.

DT 01-MAY-1997 (TREMBLREL. 03, LAST SEQUENCE UPDATE)

DT 01-MAY-1997 (TREMBLREL. 03, LAST SEQUENCE UPDATE)

DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)

DE DELTAD TRANSMEMBRANE PROTEIN PRECURSOR.

GN DELTAD.

OS BRACHYDANIO RERIO (ZEBRAFISH) (ZEBRA DANIO).

OC EUKARYOTA: METAZOA: CHORDATA: VERTEBRATA: PISCES: GNATHOSTOMATA:

OC OSTELCHTHES: ACTINOPTERYGII: CYPRINIFORMES.

RN [1]

RP SEQUENCE FROM N.A.

RA DORNESEIER P., TAKKE C., CAMPOS-ORTEGA J.A.;

RL SUBMITTED (MAR-1997) TO EMBL/GENBANK/DBJ DATA BANKS.

DR EMBL: Y11760; E307461; -

DR PROSITE: PS01187; EGF_CA; 2.

FW SIGNAL; TRANSMEMBRANE; GLYCOPROTEIN; EGF-LIKE DOMAIN.

KW SIGNAL; 19 POTENTIAL.

FT CHAIN 20 717 DELTAD TRANSMEMBRANE PROTEIN.

SO SEQUENCE 717 AA; 79061 MW; 5CC32ECA CRC32;

Query Match 30.8%; Score 1897; DB 12; Length 717;
Best Local Similarity 47.7%; Pred. No. 0.0e+00;
Matches 251; Conservative 96; Mismatches 157; Indels 22; Gaps 18;

Db 5 MIVAVLC-VNISGFCGSGVPELKLQEFVNNKGVNNGNCKG-SABEGHOC--DECKTFR 60

Qy 7 LITAFICFTYIVYVHSSGSEFLRLKTYSDHGRNDEGRCCSGSDGATGKCLGSKTR 66

Db 61 ICLKHYOANSPDPCTYGGAVTVPLGNSFQVPS--FPDSFTNPPIPAFGFTWPGT 118

Qy 67 VCLKHQATIDTISQCTYGVITPILIGENSVNLDAQFQKFTNPIDPFSSWPGT 126

Db 119 SLITELALHTDSTDLSTENPRLISRMTOHLLTVGEEMSDOLVGGRTKLKTYRFLV 178

Qy 127 SLIVEAMH-DTNNSGNARTKLLIQRLLVQVLEVSSEMTKNSQYTSLEYDFRYCD 185

Db 179 EHYEGSCAFRCPRDRTFGHFCGEGELICNSGMKGQVTEPLICLPGCDEHGFCDX 238

Qy 186 LNYGSGCAFCRPRDRTFGHFCGEGELICNSGMKGQVTEPLICLPGCDEHGFCDX 243

Db 239 GECKRCRFGFGKYDDCIRYRPGCLHGTGCOQPMOCNCOEGMGLFCNODLNTCTHHRPC 298

Qy 244 NQCVQGLGKMGALCNCEVLEPNCIHGTCKNPKWICICNMGGLXCNODLNTCTHHRPC 303

Db 299 GATCTNTGGSYTCSCRPFTGDSCEIEVNEC-SG-SPCRNAGSGCTD-LE-NT-YSCTCP 353

Qy 304 GGCFTNTEGELYTCCKAPYSGSDCENEIYSCDADVPNCPQNGGTCIDEPHTKTYGKCH 363

Db 354 PGFTYGRNCELAMTCAAGPCPNNG--HC-AD-NPEG-GYFCQCMGAGNCEKIDHCS 408

Qy 364 NGMSGKCEKRYLVCSKPCRHOGICRNVVRGLSGKGGYQCEPCIGYSGPCNCDQLDNCS 423

Db 409 SNFCSNADCLDLDVSLGCGPFGTGHCEIDNDECATYPCONGGCTDLYTCTCP 468

Qy 424 PNCINGSGC-QPSGK--CICPAGFSGTRETNTIDDLGHCENGGTCTIDMNVQRCV 480

Db 469 PGYTKNCTSAVNKCLNPHNGATCEHMDNRYACACIPGYGRNC 514

Qy 481 PGHGHGHSKVDCLIRPCANGTCLNNDYQCTCRAGFTGKDCSVYDIDECSSGP 526

DE DELTA.

OS HOMO SAPIENS (HUMAN).

OC EUKARYOTA: METAZOA: CHORDATA: VERTEBRATA: TETRAPODA: MAMMALIA:

OC EUTHERIA: PRIMATES.

RN [1]

RP SEQUENCE FROM N.A.

RA MANN R.S., GRAY G.E., HENRIQUE D., ISH-HOROWICZ D.,

RL SUBMITTED (JUN-1997) TO EMBL/GENBANK/DBJ DATA BANKS.

DR EMBL: AF003532; G2197069; -

DR PROSITE: PS01187; EGF_CA; 1.

FW GLYCOPROTEIN; EGF-LIKE DOMAIN.

SO SEQUENCE 723 AA; 77956 MW; A1D48BDB CRC32;

Query Match 30.3%; Score 1865; DB 2; Length 723;
Best Local Similarity 48.4%; Pred. No. 0.0e+00;
Matches 261; Conservative 86; Mismatches 171; Indels 21; Gaps 17;

Db 4 RCALALVALSLCQVMSGVPELKLQEFVNNKGLGNRNCRCGA-GPP-C--ACRTF 59

Qy 5 KCLTAFICFTYIVYVHSSGSEFLRLKTYSDHGRNDEGRCCSGSDGATGKCLGSKTR 64

Db 60 FRVCLKHQASVPEPCTYGGAVTVPLGVSFSLPDGG-ADSAFNPPIRPFPGTWP 118

Qy 65 FRVCLKHQATIDTISQCTYGVITPILIGENSVNLDAQFQKFTNPIDPFSSWPG 124

Db 119 TFSLITELALHTDSTDLSTENPRLISRMTOHLLTVGEEMSDOLVGGRTKLKTYRFLV 178

Qy 125 TFSLIVEAMH-DTNNSGNARTKLLIQRLLVQVLEVSSEMTKNSQYTSLEYDFRYT 183

Db 179 CDEHYEGSCAFRCPRDRTFGHFCGEGELICNSGMKGQVTEPLICLPGCDEHGFCD 238

Qy 184 CDLNTYSGCAFCRPRDRTFGHFCGEGELICNSGMKGQVTEPLICLPGCDEHGFCD 241

Db 239 KPECKRCRFGFGKYDDCIRYRPGCLHGTGCOQPMOCNCOEGMGLFCNODLNTCTHHRPC 298

Qy 242 KPNQCVQGLGKMGALCNCEVLEPNCIHGTCKNPKWICICNMGGLXCNODLNTCTHHRPC 301

Db 299 KNGATCTNTGGSYTCSCRPFTGDSCEIEVNEC-SG-SPCRNAGSGCTD-LE-NT-YSCTCP 353

Qy 302 KNGCFTNTEGELYTCCKAPYSGSDCENEIYSCDADVPNCPQNGGTCIDEPHTKTYGKCH 361

Db 354 CPGFTYGRNCELAMTCAAGPCPNNG--HC-AD-NPEG-GYFCQCMGAGNCEKIDHCS 408

Qy 362 CANGSGKCEKRYLVCSKPCRHOGICRNVVRGLSGKGGYQCEPCIGYSGPCNCDQLDN 421

Db 409 CSSPFCSGKACVYDGLDYLRCQAGFSGRHCDVNDVDCASSPCANGTCTRQVNDPFGCT 468

Qy 422 CSPNDCINGSGCQPSGK--CICPAGFSGTRETNTIDDLGHCENGGTCTIDMNVQRCV 478

Db 469 CPGFTYGRNCELAMTCAAGPCPNNG--HC-AD-NPEG-GYFCQCMGAGNCEKIDHCS 514

Qy 479 CVPGHGHGHSKVDCLIRPCANGTCLNNDYQCTCRAGFTGKDCSVYDIDECSSGP 527

RESULT 5 PRELIMINARY; PRT; 721 AA.

ID 091902; PRELIMINARY; PRT; 721 AA.

AC 091902; PRELIMINARY; PRT; 721 AA.

DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)

DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)

DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)

DE X-DELTA-1.

OS XENOPUS LAEVIS (AFRICAN CLAMMED FROG).

OC EUKARYOTA: METAZOA: CHORDATA: VERTEBRATA: TETRAPODA: AMPHIBIA: ANURA.

RN [1]

RP SEQUENCE FROM N.A.

RA MEDLINE: 95319507.

RA HENRIQUE D., ADAM J., CHITNIS A., LEWIS J., ISH-HOROWICZ D.,

RL NATURE 375:787-790(1995).

RN [2]

SO SEQUENCE FROM N.A.

RX MEDLINE: 95319503.

RX CHITNIS A.B., HENRIQUE D., LEWIS J., ISH-HOROWICZ D., KINTNER C.R.;

QY 7 LTTAFICFTVIYOVH--SSGSFELRLKYPFNDHGRDNEGRCCSGESDGTGKCL--GSKCTR 64
Db 75 FVCLKHYQATIDTTSOCTGYDVITPILIGENSVNLTDAORFNKGFNTNIOFPFESFWPG 130
QY 65 FVCLKHYQATIDTTSOCTGYDVITPILIGENSVNLTDAORFNKGFNTNIOFPFESFWPG 124
Db 131 STLLVYEA--DSSND--TVOPDSI--IEKASHSGMINPSROMOTLKONTGVAHEYOIRVTC 187
QY 125 TFSLVEAHMDTNNNGNARTNKLRLIOLRLVQOVLVEVSESEKTKNSQYTSLEYDERVTC 184
Db 188 DYYYGFGCNKCRPRDDFFGHYACDQNGNKTCEMGMWPECNRAICRGSGPKHSGCL 247
QY 185 DLNYYGSGCAKFCRPRDSSFHSTGSETEGELICLTGMODYCHIPKANGCE--HGHCOK 242
Db 248 PGDCRCOYGMOLGYDKCIPHPGCVHIGICNEPMOCICETNMGGOLCDKDLNCTGTHQPC 307
QY 243 PNOVCYQOLGMKALCNECVLEPNICHTGNKRPWTCICNEGMGLYCNODLNTCTNHRPK 302
Db 308 NGGTCSTNGPDKYOCSCPEYSGPNCIEAHACLSD--PCNHRGSC--KE--TSLGFECEC 362
QY 303 NGGTCFNTGEGLYTCKCARGYSGDCENEIYSCDADVNPCONGGTCIDEPHTKTYKCHC 362
Db 363 SPGWTGPTCSTNIDDCSPNMCNCSHGTCOD---L-VNG--FKCVPPOMTGTQOLDANE 415
QY 363 ANGMSGKMCERKVLTCSDKPC--HOGICRNVRPGLSGKGGYOCCEPIGYSGPCDLOLDN 421
Db 416 CEAKPCVNAKSCKNLIASYYCDLPGWNGONCDININDCLG--OCONDASCRDLVNGYRCI 474
QY 422 CSPNFCINGSCOP--SGK--CICPAGFSGTRETINIDCLGHCENGCTCIDMVMQYRCQ 478
Db 475 CPBGYAGDHCEIRDIECASNPCLNGHCONEINRQCICPTGSGNLCOLDIDYCEPNPC 534
QY 479 CVPGFHGTSSKVDCLIRPCANGGTCLNLNDYQCTCRAGFTGDKDSVDDIECSSGPC 538
Db 535 ONGAOCYNRASDYFCPCPEDEYKNC 560
QY 539 HNGGTGMNRVNSFEVCANGFRGKOC 564

RESULT 8
ID 014902 PRELIMINARY; PRT; 1218 AA.
AC 014902:
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DE 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE TRANSMEMBRANE PROTEIN JAGGED 1.
GN H21.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 95211842.
RA LINDELL C.E., SHAMBER C.J., BOULTER J., WEINMASTER G.;
RL CELL 80:909-917(1995).
RN [2]
RP SEQUENCE FROM N.A.
RA BASH J., ZONG W.-X., GELINAS C.;
RL SUBMITTED (NOV-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: AF028593; G2599082; -
DR PROSITE: PS01187; EGF_CA; 8.
KW TRANSMEMBRANE; GLYCOPROTEIN; EGF-LIKE DOMAIN.
SQ SEQUENCE 1218 AA; 133798 MW; 07B97EE3 CRC32;
Query Match 27.4%; Score 1691; DB 2; Length 1218;
Best Local Similarity 41.7%; Pred. No. 0.00e+00;
Matches 236; Conservative 109; Mismatches 192; Indels 29; Gaps 19;
Db 16 LLLALIC-ALRAKYGASGOFELEILSMQNVNGELONGCGGARNPDRCRTRECDTY 74
QY 7 LTTAFICFTVIYOVH--SSGSFELRLKYPFNDHGRDNEGRCCSGESDGTGKCL--GSKCTR 64
Db 75 FVCLKHYQATIDTTSOCTGYDVITPILIGENSVNLTDAORFNKGFNTNIOFPFESFWPG 130

QY 65 FVCLKHYQATIDTTSOCTGYDVITPILIGENSVNLTDAORFNKGFNTNIOFPFESFWPG 124
Db 131 STLLVYEA--DSSND--TVOPDSI--IEKASHSGMINPSROMOTLKONTGVAHEYOIRVTC 187
QY 125 TFSLVEAHMDTNNNGNARTNKLRLIOLRLVQOVLVEVSESEKTKNSQYTSLEYDERVTC 184
Db 188 DYYYGFGCNKCRPRDDFFGHYACDQNGNKTCEMGMWPECNRAICRGSGPKHSGCL 247
QY 185 DLNYYGSGCAKFCRPRDSSFHSTGSETEGELICLTGMODYCHIPKANGCE--HGHCOK 242
Db 248 PGDCRCOYGMOLGYDKCIPHPGCVHIGICNEPMOCICETNMGGOLCDKDLNCTGTHQPC 307
QY 243 PNOVCYQOLGMKALCNECVLEPNICHTGNKRPWTCICNEGMGLYCNODLNTCTNHRPK 302
Db 308 NGGTCSTNGPDKYOCSCPEYSGPNCIEAHACLSD--PCNHRGSC--KE--TSLGFECEC 362
QY 303 NGGTCFNTGEGLYTCKCARGYSGDCENEIYSCDADVNPCONGGTCIDEPHTKTYKCHC 362
Db 363 SPGWTGPTCSTNIDDCSPNMCNCSHGTCOD---L-VNG--FKCVPPOMTGTQOLDANE 415
QY 363 ANGMSGKMCERKVLTCSDKPC--HOGICRNVRPGLSGKGGYOCCEPIGYSGPCDLOLDN 421
Db 416 CEAKPCVNAKSCKNLIASYYCDLPGWNGONCDININDCLG--OCONDASCRDLVNGYRCI 474
QY 422 CSPNFCINGSCOP--SGK--CICPAGFSGTRETINIDCLGHCENGCTCIDMVMQYRCQ 478
Db 475 CPBGYAGDHCEIRDIECASNPCLNGHCONEINRQCICPTGSGNLCOLDIDYCEPNPC 534
QY 479 CVPGFHGTSSKVDCLIRPCANGGTCLNLNDYQCTCRAGFTGDKDSVDDIECSSGPC 538
Db 535 ONGAOCYNRASDYFCPCPEDEYKNC 560
QY 539 HNGGTGMNRVNSFEVCANGFRGKOC 564
RESULT 9
ID 015122 PRELIMINARY; PRT; 1218 AA.
AC 015122:
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DE 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE JAGGED1.
GN JAG1.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
RP SEQUENCE FROM N.A.
RA ODA T., ELKAHLOUN A.G., PIRE B.L., OKAJIMA K., KRANTZ I.D., GENIN A.,
RA PICCOLI D.A., MELTZER P.S., SPINNER N.B., COLLINS F.S.,
RA CHANDRASEKHARAPPA S.C.;
RL NAT. GENET. 0:0-0(1997).
RN [2]
RP SEQUENCE FROM N.A.
RA ODA T., ELKAHLOUN A.G., MELTZER P.S., CHANDRASEKHARAPPA S.C.;
RL GENOMICS 43:376-379(1997).
DR EMBL: AF003837; G2228793; -
DR PROSITE: PS01187; EGF_CA; 8.
KW GLYCOPROTEIN; EGF-LIKE DOMAIN.
SQ SEQUENCE 1218 AA; 133858 MW; 20F471DB CRC32;
Query Match 27.4%; Score 1687; DB 2; Length 1218;
Best Local Similarity 41.5%; Pred. No. 0.00e+00;
Matches 235; Conservative 110; Mismatches 192; Indels 29; Gaps 19;
Db 16 LLLALIC-ALRAKYGASGOFELEILSMQNVNGELONGCGGARNPDRCRTRECDTY 74
QY 7 LTTAFICFTVIYOVH--SSGSFELRLKYPFNDHGRDNEGRCCSGESDGTGKCL--GSKCTR 64
Db 75 FVCLKHYQATIDTTSOCTGYDVITPILIGENSVNLTDAORFNKGFNTNIOFPFESFWPG 130
QY 65 FVCLKHYQATIDTTSOCTGYDVITPILIGENSVNLTDAORFNKGFNTNIOFPFESFWPG 124

```
Db 131 SYTLLEAV--DSSND--TVQPSDI--IEKASHGMINSPRQHOTLKQNTGVAHFYQIRVTC 187
QY 125 TFSLLVEAWHDTNNSGNARTKLLIRLLVQVLEVSSEKTKNSQYTSLEYDFRVT 184
Db 188 DDYYGFCNKFCRPRDDFGHYACDQNGKNTCMGWMGRCNRAICROGCKSPKHGSKL 247
QY 185 DLNYGSGCAKFCRPRDDSGHSTCTSETGEIICLTGHOQDYCHIPCAKGE--HGCHDK 242
Db 248 PGDCRCQYGHQGLYCDKCIPIHPCVGHGTCTNEPWCICETNWGGOLCDKDLNYCGTHQPCL 307
QY 243 PNCVCOLGKAGALCNVLEPCNIHGTCNKPTWCICNEGWWGLYCNOGLNYCTNHRPCK 302
Db 308 NGGTCNTGTPDKYQCSPEGYSPNCEIAEHACLS--PCNHRGSC--KE--TSSGFCEC 362
QY 303 NGGTCNTGEGLYTCAPGSGDDCENEIYSCDADVPNCQNGTCTDEPHTKTYGKCHC 362
Db 363 SPGWGTGPTCTNIDDCSPNNCSHGTCQD---L-VNG---FKVCPPQWTGKTCOLDANE 415
QY 363 ANGSWKMEKVKLTCDKPC-HQICRNVPRGLSGQGYQCEPCIGYSGPNCDLQDN 421
Db 416 CEAKPCVNAKSKNLIAYYCDCLPGWMQNCININDCLG-QCQNDASCRDLVNGYRCI 474
QY 422 CSPNCPINGSCQP-SGK--CICPAGSGTRCETNIDDCLGHCENGCTCIDWVNYRQC 478
Db 475 CPPGYAGDHCEERDIDECASNPCLDGGHCONEINRFQCLPTGFGSGLCQDIDYCEPNPC 534
QY 479 CVPFGHGTCHSSKVDLIRPCANGGTCLNNDYQCTCRAGFTGKDCSVDIDECSSGPC 538
Db 535 ONGAQCYNRADYFCKCPEDYEGKNC 560
QY 539 HNGGTCHNRVNSFEVCANGFRGKQC 564

RESULT 10
ID Q63722 PRELIMINARY; PRT; 1219 AA.
AC Q63722; P70640;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-FEB-1997 (TREMBLREL. 02, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE JAGGED PROTEIN.
OS RATTUS NORVEGICUS (RAT).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SCIATIC NERVE;
RX MEDLINE; 95211842.
RA LINDSELL C.E., SHAWBER C.J., BOULTER J., WEINMASTER G.;
RL CELL 80:909-917(1995).
DR EMBL; L38483; G1492111; -.
DR PROSITE; PS01187; EGF_CA; 8.
KW GLYCOPROTEIN; EGF-LIKE DOMAIN.
SQ SEQUENCE 1219 AA; 134325 MW; B193F948 CRC32;

Query Match 27.4%; Score 1692; DB 10; Length 1219;
Best Local Similarity 41.7%; Pred. No. 0.00e+00;
Matches 236; Conservative 111; Mismatches 190; Indels 29; Gaps 19;

Db 16 LLLALIC-ALRAKYCGASGQFEILSMQNVNGLQNCGCCGARNPGDRCKTRDECDTY 74
QY 7 LTTAFICFTVIVQVH--SSGSFELRLKYFSNDHGRDNEGRCCSGESDGTGKCL-GSKCTR 64
Db 75 FKVCLKEYQSRVTAGGCSFGSGSTPYGGNTFNKLAS-RGNDR--NRVLFPFSFAWRP 130
QY 65 FRVCLKHQATIDTTSQCTGVDVITPILGENSVNLTAQRPQNGFTNPQTFPFSFSPG 124
Db 131 SYTLLEAV--DSSND--TVQPSDI--IEKASHGMINSPRQHOTLKQNTGVAHFYQIRVTC 187
QY 125 TFSLLVEAWHDTNNSGNARTKLLIRLLVQVLEVSSEKTKNSQYTSLEYDFRVT 184
Db 188 DDYYGFCNKFCRPRDDFGHYACDQNGKNTCMGWMGRCNRAICROGCKSPKHGSKL 247
QY 185 DLNYGSGCAKFCRPRDDSGHSTCTSETGEIICLTGHOQDYCHIPCAKGE--HGCHDK 242
```

```
QY 185 DLNYGSGCAKFCRPRDDSGHSTCTSETGEIICLTGHOQDYCHIPCAKGE--HGCHDK 242
Db 248 PGDCRCQYGHQGLYCDKCIPIHPCVGHGTCTNEPWCICETNWGGOLCDKDLNYCGTHQPCL 307
QY 243 PNCVCOLGKAGALCNVLEPCNIHGTCNKPTWCICNEGWWGLYCNOGLNYCTNHRPCK 302
Db 308 NGGTCNTGTPDKYQCSPEGYSPNCEIAEHACLS--PCNHRGSC--KE--TSSGFCEC 362
QY 303 NGGTCNTGEGLYTCAPGSGDDCENEIYSCDADVPNCQNGTCTDEPHTKTYGKCHC 362
Db 363 SPGWGTGPTCTNIDDCSPNNCSHGTCQD---L-VNG---FKVCPPQWTGKTCOLDANE 415
QY 363 ANGSWKMEKVKLTCDKPC-HQICRNVPRGLSGQGYQCEPCIGYSGPNCDLQDN 421
Db 416 CEAKPCVNAKSKNLIAYYCDCLPGWMQNCININDCLG-QCQNDASCRDLVNGYRCI 474
QY 422 CSPNCPINGSCQP-SGK--CICPAGSGTRCETNIDDCLGHCENGCTCIDWVNYRQC 478
Db 475 CPPGYAGDHCEERDIDECASNPCLDGGHCONEINRFQCLPTGFGSGLCQDIDYCEPNPC 534
QY 479 CVPFGHGTCHSSKVDLIRPCANGGTCLNNDYQCTCRAGFTGKDCSVDIDECSSGPC 538
Db 535 ONGAQCYNRADYFCKCPEDYEGKNC 560
QY 539 HNGGTCHNRVNSFEVCANGFRGKQC 564

RESULT 11
ID P78504 PRELIMINARY; PRT; 1227 AA.
AC P78504;
DT 01-MAY-1997 (TREMBLREL. 03, CREATED)
DT 01-MAY-1997 (TREMBLREL. 03, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE JAGGED 1 (TRANSMEMBRANE PROTEIN JAGGED).
GN HJ1.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 95211842.
RA LINDSELL C.E., SHAWBER C.J., BOULTER J., WEINMASTER G.;
RL CELL 80:909-917(1995).
RN [2]
RP SEQUENCE FROM N.A.
RP LI L., DENG Y., BANTA A.B., HOOD L.;
RA SUBMITTED (DEC-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [3]
RP SEQUENCE OF 14-1227 FROM N.A.
RX MEDLINE; 97115768.
RA ZIMRIN A.B., PEPPER M.S., MCMAHON G., NGUYEN F., MONTESANO R.,
RA MACIAG T.;
RL J. BIOL. CHEM. 271:32499-32502(1996).
RN [4]
RP REVISIONS TO 14-1227.
RA ZIMRIN A.B., NGUYEN F., MACIAG T.;
RL SUBMITTED (MAY-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; U73936; G1695274; -.
DR EMBL; U7720; E319096; -.
KW TRANSMEMBRANE.
FT CONFLICT 1187 1227
FT -> NGTPTKHPNTWKQNDRLDLESAQSLNRMEIY
FT (IN REF. 1 AND 2).
SQ SEQUENCE 1227 AA; 134770 MW; 5D300881 CRC32;

Query Match 27.4%; Score 1691; DB 2; Length 1227;
Best Local Similarity 41.7%; Pred. No. 0.00e+00;
Matches 236; Conservative 109; Mismatches 192; Indels 29; Gaps 19;

Db 16 LLLALIC-ALRAKYCGASGQFEILSMQNVNGLQNCGCCGARNPGDRCKTRDECDTY 74
QY 7 LTTAFICFTVIVQVH--SSGSFELRLKYFSNDHGRDNEGRCCSGESDGTGKCL-GSKCTR 64
```


Db 353 AIDIDECASNPAGGCTC-ID-H--INS-----FECICPQWIGATCQLDANECEGKPCVNA 405
QY EVKLTSCDKPCHQG-ICRNVPRGLSGKGGYQCECPIGYSGPNCDLQLDNCSNPNPCING 430
Db 406 YSKNLIIGGYCDIPGWGVNCHINIDCHG-QCOHGGTCKDEVNDYHCICPRGFTGKN 464
QY 431 GSCQ--SGK-CICPAGFSGTRCTETIDDCGLHQCENGGTCIDMNYRCQCPVGFHGT 487
Db 465 CEIETNECESNPQNGCRKDLNGFTCLCAQFGSGVFCMDIDFCEPNPCQNGAKCYDL 524
QY 488 CSSKVDLCLIRPCANGTCLNLNDYQCTCRAGFTGKDCSVDIDECSSGCHNGGTGCMNR 547
Db 525 GDDYACAPDDYDKNC 541
QY 548 VNSFECVCANGFRGKC 564

RESULT 14
ID P97607; PRELIMINARY; PRT; 1202 AA.
AC P97607;
DT 01-MAY-1997 (TREMBLREL. 03, CREATED)
DT 01-MAY-1997 (TREMBLREL. 03, LAST SEQUENCE UPDATE)
DE 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE JAGGED2 (FRAGMENT)
OS RATTUS NORVEGICUS (RAT).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A.
RA SHAWBER C.J., BOULTER J., LINDSELL C.E., WEINMASTER G.;
RL DEV. BIOL. 180:370-376(1996).
DR EMBL; U70050; G1718248; -;
DR PROSITE; PS01187; EGF_CA; 7.
KW GLYCOPROTEIN; EGF-LIKE DOMAIN.
FT NON_TER 1
SQ SEQUENCE 1202 AA; 129704 MW; 697F4205 CRC32;

Query Match 24.5%; Score 1510; DB 10; Length 1202;
Best Local Similarity 41.6%; Pred. No. 0.00e+00;
Matches 225; Conservative 96; Mismatches 186; Indels 34; Gaps 22;

Db 1 GACCDGTRTRTGGCGRDCDIYVRVCLKEYOAKYTPPTGCSYGYGATPVLGNSYFLP 60
QY 43 GRCCGESDGA-TGKC-LGSKTRFRVCLKHQYQATIDTTSQCTYGDVITPILGENSVNL- 99
Db 61 PAGAAGDRARARTRGGHODPGLVVPFQPAWPRSFLLIVEAW-DWMDND-TTPDEELLIE 118
QY 100 -TDA---Q-RFQNK--GFTNP--IQPFPSFSPGTFSLIVEAWHDTNNSGNARTKLLIQ 150
Db 119 RVSHAGMINPEDRWKSLHFSGHVAHLELQIRVRCDENYYSATCNKFCRPRNDFGHVYTC 178
QY 151 RLIVQOVLEVSVSEWKTNKSESYTSLEYDPRVTCDLNYYSGGCAKFCRPRDDSFHSTCS 210
Db 179 QYGNKACMDGWMKECKEAKVCKOGCNLLHGGCTVPGECCRSYGNQGFCECVPYPCVH 238
QY 211 ETGEIICLTGWOQDCHIPKCAKCE--HGCDKPNQCVQQLGKALCNCEVLEPNCH 268
Db 239 GSCVEPHWCDENWGLLCKDLNLYCGSHHPCVNGGTCINAEPDQYLCACPDGYLGKNC 298
QY 269 GTCNKPTCICNKGWGLYCNQDLNLYCTNHRPKNGGTCFTNTEGLYTCCKAPGYSDDC 328
Db 299 ERAEHAC-AS-NPCANGSC-HE--VLSGFECCHPCPSWGTPCALDIDECASNPNCAAGGT 353
QY 329 ENEIYSCDADVPNCGGTCTIDEPHTKTGYKCHCANGSGKMCCKEYVLTCSDKPCHQG-I 387
Db 354 C--VDQ-V----DGFECICPEQWVGATCQLDANECEGKPCPLNAPFCKNLIGGIYCDCLPG 406
QY 388 CRNVPRGLSGKGGYQCECPIGYSGPNCDLQLDNCSNPNPCINGSCQP--SGK-CICPAG 444
Db 407 WKGANCHINIDCHG-QCOHGGTCKDLVNGYQCVPRGFGGRHCELEYKCASSPCRGG 465
QY 445 FSGTRCTETIDDCGLHQCENGGTCIDMNYRCQCPVGFHGTGCHCSKVDLCLIRPCANGG 504

Db 466 ICEDLYDGFRCPRGLSGPLCEVDVLDWCEPNPLNGARCYNLEDDYACACPEDFGGKN 525
QY 505 TCLNLNDYQCTCRAGFTGKDCSVDIDE-CSSGCHNGGTGCMNRVNSFECVCANGFRGKC 563
Db 526 C 526
QY 564 C 564

RESULT 15
ID Q06008; PRELIMINARY; PRT; 1203 AA.
AC Q06008;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE NOTCH2 OR NOTCH B.
DE NOTCH2 OR NOTCH B.
GN NOTCH2 OR NOTCH B.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-F1 (CBA X C57BL); TISSUE-WHOLE EMBRYO;
RX MEDLINE; 93178563
RA LARDELLI M., LENDAHU U.;
RL EXP. CELL RES. 204:364-372(1993).
DR EMBL; X68279; G287990; -;
DR MGD; MGI:97364; NOTCH2.
KW DIFFERENTIATION; NEUROGENESIS; REPEAT.
FT NON_TER 1
FT NON_TER 1203
SQ SEQUENCE 1203 AA; 128982 MW; ASA95551 CRC32;

Query Match 15.7%; Score 970; DB 10; Length 1203;
Best Local Similarity 40.9%; Pred. No. 2.11e+195;
Matches 146; Conservative 74; Mismatches 108; Indels 29; Gaps 17;

Db 401 YSQVNECLSNPCHNGTGLSGYKCLCDAGWGVNCEVDKNECLSNP-CONGTGTCNNLV 459
QY 225 YCHIPKC-ARGCEHGHG-DK-PN-QCYCQLGWKGALC----NECVLEPNCHIG-TCNK-- 273
Db 460 NGYRCTCKGFKGYNCQVNDIECASN-PCLNQGTCTFDDVSG-YTCHCMPLPYTKNQCOTVL 517
QY 274 -PWTICNEGWGLYCNODLNYCTNHRPKNGGTCTNTEGLYTCCKAPGYSGGDCENEI 332
Db 518 APCSP--NPCENAAVCKEAPNFES-FSLCAPGQGRCTVDVDECISKPCMNNGVCHNT 574
QY 333 YSCDADYNPCQNGGTCIDEPHTKTGYKCHCANGSGKMCCKEYVLTCSDKPC-HQIGICNV 391
Db 575 Q---GS-----YVCECPGFGSGMDCDEEDINDCLANPCQNGSCVDHYNTFSCQHPGFIGD 627
QY 392 RPLGLSGKGGYQCECPIGYSGPNCDLQLDNCSNPNPCINGGSC--QPSG-KCICPAGFSGT 448
Db 628 KCQTDMECLSEPCCKNGGTCSDYVNSYTCCTCPAGFHGVHCENNIDECESSCFNGGTCVD 687
QY 449 RCTETIDDCGLHQCENGGTCIDMNYRCQCPVGFHGTGCHCSKVDLCLIRPCANGGTCNLN 508
Db 688 GINSFSLCPVGFPGFCLHDIENECSSNPCLNAGTCVDGLGTGYRCICPLGYTGKNQ 744
QY 509 LNNDYQCTCRAGFTGKDCSVDIDECSSGCHNGGTGCMNRVNSFECVCANGFRGKCQD 565

Search completed: Fri Jun 12 11:58:15 1998
Job time : 163 secs.


```

Db 181 rvtcdlnyvggcafcprddsfghstcsetgeiicltgwgdychipk 230
QY 181 RVTCDLNYGSGCAFCRPRDDSFHSTCSETGEIICLTGWOGDYCHIPK 230

RESULT 2
ID R28960 standard; Protein; 833 AA.
AC R28960,1993 (first entry)
DT 01-APR-1993
DE Delta Dll.
KW Human; Notch; plasmid; cDNA; clone; Dll; expression library; PCR;
KW polymerase chain reaction; primer; cloning vector; Delta; Serrate;
KW neurogenic; topotypic; homotypic; heterotypic; differentiation;
KW quantitation; antibody.
OS Homo sapiens.
PN W09219734-A.
PD 12-NOV-1992.
PF 01-MAY-1992; U03651.
PR 03-MAY-1991; US-695189.
PR 14-NOV-1991; US-791923.
PA (INDV ) UNIV INDIANA FOUND.
PA (UYIA ) UNIV YALE.
PI Artavanis-Tsakonas S, Blaumueller CM, Fehon RG, Muskavitch MAT;
PI Rebay I, Shepard SB;
DR WPI: 92-398861/48.
DR N-PSDB: Q30997.
PT Human Notch and Delta DNA and protein sequences - used for study
PT and manipulation of differentiation processes
PS Claim 50; Fig 13; 239pp; English.
CC The sequence given is encoded by the nucleotide sequence of human
CC Delta gene contained in plasmid cDNA clone Dll. A human expression
CC library was constructed and screening assays were carried out on to
CC select for the expressed Delta product. Alternatively the sequences
CC could be isolated by amplification using polymerase chain reaction
CC (PCR) primers. The isolated gene may be inserted into a cloning
CC vector and expressed. The Delta gene and also the Notch and Serrate
CC neurogenic genes are designated "topotypic" genes. The proteins
CC they encode are involved in specific homo- or heterotypic interactions
CC crucial to differentiation. The quantitation of mRNA for human Notch
CC and Delta and adhesive molecules, and study of its expression are
CC possible using the DNA and antibodies raised against the Notch and
CC Delta proteins.
SQ Sequence 833 AA;

Query Match 100.0%; Score 1734; DB 6; Length 833;
Best Local Similarity 100.0%; Pred. No. 2.83e-176;
Matches 230; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 mhwkclltaficftvqvhsqsfelrkyfndhgrdnegrcsgsdgatkclgs 60
QY 1 MHWIKCLLTAFICFTVIVQVHSSGSFELRLKYFSDHGRDNEGRCSCGESDGTGKCLGS 60

Db 61 cktfrvclkhqyatidtsqctygvitpiligensvnltdaqrfnkgfntpiqpfsf 120
QY 61 CKTRFVCLKHQYATIDTSQCTYGVITPILIGENSVNLTDQRFNKGFTNPQPFPSF 120

Db 121 swpgtflsiveahwdnnsgnartnklllqrlvqvlevssekwnksesqytsleydf 180
QY 121 SWPGTFLSIVEAHDNNSGNARTNKKLLIQLRLVQVLEVSEKWNKSESYTSLEYDF 180

Db 181 rvtcdlnyvggcafcprddsfghstcsetgeiicltgwgdychipk 230
QY 181 RVTCDLNYGSGCAFCRPRDDSFHSTCSETGEIICLTGWOGDYCHIPK 230

RESULT 3
ID W11719 standard; Protein; 727 AA.
AC W11719;
DT 28-APR-1997 (first entry)
DE C-Delta-1 polypeptide.
KW C-Delta-1; cell proliferation; nervous system disorder;
KW tissue regeneration; Notch; cervix cancer; breast cancer;

```

```

KW lung cancer; colon cancer; melanoma; seminoma;
KW neurogenesis; therapy.
OS Gallus sp.
FH Key Location/Qualifiers
FT domain 184..228
FT /label= DSL
FT domain 229..261
FT /label= EGF1
FT domain 262..292
FT /label= EGF2
FT domain 293..332
FT /label= EGF3
FT domain 333..370
FT /label= EGF4
FT domain 371..409
FT /label= EGF5
FT domain 410..447
FT /label= EGF6
FT domain 448..485
FT /label= EGF7
FT domain 486..523
FT /label= EGF8
FT domain 524..534
FT /label= EGF9
FT domain 555..579
FT /label= TM
FT /note= "transmembrane domain"
PN W09701571-A1.
PD 16-JAN-1997.
PF 28-JUN-1996; U11178.
PR 28-JUN-1995; US-000589.
PA (IMCR ) IMPERIAL CANCER RES TECHNOLOGY.
PA (UYIA ) UNIV YALE.
PI Artavanis-Tsakonas S, Gray GE, Henrique D, Ish-Horowicz D;
PI Lewis J;
DR WPI: 97-100159/09.
DR N-PSDB: T58897.
PT New vertebrate Delta protein, DNA and antibodies - for treating and
PT preventing cancer, nervous system disorders and for tissue
PT regeneration
PS Disclosure; Fig 2; 135pp; English.
CC C-Delta-1 polypeptide (W11719) is the chick homologue of Drosophila
CC Delta, a protein that binds to Notch protein. Expression of
CC C-Delta-1 correlates with onset of neurogenesis. The C-delta-1
CC amino acid sequence was deduced from a cDNA clone (T58897) obtd.
CC from chick stage 4-6 embryos. An alternatively spliced variant
CC (W00876) was also isolated, and mouse (W11720) and human (W11721-
CC 38) Delta-1 polypeptides have been identified. Delta-1 proteins
CC can be used to treat or prevent disorders characterised by
CC increased Notch activity, such as cervical, breast, lung or colon
CC cancer, melanoma or seminoma, and nervous system disorders or to
CC promote tissue regeneration and repair.
SQ Sequence 727 AA;

Query Match 44.7%; Score 775; DB 21; Length 727;
Best Local Similarity 49.1%; Pred. No. 2.25e-70;
Matches 104; Conservative 33; Mismatches 74; Indels 1; Gaps 1;

Db 20 qvdsgvqfklqefvnnkglalnrcncrvgggpggagqgqcdcktfrrvclkhqvasvp 79
QY 19 QVHSSGSFELRLKYFSDHGRDNEGRCSCGESDGTGKCLGSCKTRFVCLKHQYATIDT 78

Db 80 epctvgysaitpvlansfsvpdgagadpafsnpirfpgftwptfslilialhtdsp 139
QY 79 TSQCTYGVITPILIGENSVNLTDQRFNKGFTNPQPFPSFSGPSTFLSIVEAWH-DTN 137

Db 140 ddlttenperlisrlatrhlaygeewsdqlhssgtrdtkysvrfvcdelygqgsvfc 199
QY 138 NSGNARTNKKLLIQLRLVQVLEVSEKWNKSESYTSLEYDFRVTCDLNYGSGCAKFC 197

Db 200 rprddrfghftcgcgkvcnpgwkgqyctep 231
QY 198 RPRDSDFGHSTCSETGEIICLTGWOGDYCHIP 229

```

ID	RESULT	4
AC	M00876 standard; Protein; 740 AA.	
AD	M00876:	
DE	28-APR-1997 (first entry)	
DE	C-delta-1 polypeptide (alternatively spliced variant).	
KM	C-delta-1; cell proliferation; nervous system disorder;	
KM	tissue regeneration; Notch; cervix cancer; breast cancer;	
KM	lung cancer; colon cancer; melanoma; seminoma;	
KM	neurogenesis; therapy.	
OS	Gallus sp.	
EH	Key	
EH	Location/Qualifiers	
EH	domain	184..228
EH	domain	/label= DSL
EH	domain	229..261
EH	domain	/label= EGF1
EH	domain	262..292
EH	domain	/label= EGF2
EH	domain	293..332
EH	domain	/label= EGF3
EH	domain	333..370
EH	domain	/label= EGF4
EH	domain	371..409
EH	domain	/label= EGF5
EH	domain	410..447
EH	domain	/label= EGF6
EH	domain	448..485
EH	domain	/label= EGF7
EH	domain	486..523
EH	domain	/label= EGF8
EH	domain	524..534
EH	domain	/label= EGF9
EH	domain	535..579
EH	domain	/label= TM
EH	domain	/note= "transmembrane domain"
PN	W09701571-A1.	
PD	16-JAN-1997.	
PF	28-JUN-1996: U11178.	
PR	28-JUN-1995: US-000589.	
PA	(IMCR) IMPERIAL CANCER RES TECHNOLOGY.	
PI	(UYVA) UNIT YALE.	
PI	Aravaniis-Tsakonas S, Gray GE, Henrique D, Ish-Horowitz D;	
PI	Lewis J;	
PI	WPI: 97-100159/09.	
DR	N-PSDB: T58898.	
DR	New vertebrate Delta protein, DNA and antibodies - for treating and	
PT	preventing cancer, nervous system disorders and for tissue	
PT	regeneration	
PS	Disclosure: Fig 2: 135pp: English.	
CC	C-delta-1 polypeptide (W00876) is the chick homologue of Drosophila	
CC	Delta, a protein that binds to Notch protein. Expression of	
CC	C-delta-1 correlates with onset of neurogenesis. The C-delta-1	
CC	amino acid sequence was deduced from a cDNA clone (U58898) obtd.	
CC	from chick stage 4-6 embryos. A shorter version (W58877) of	
CC	C-delta-1, lacking the 12 C-terminal amino acids of the longer	
CC	version, was also isolated, and mouse (W11720) and human (W11721-	
CC	38) Delta-1 polypeptides have been identified. Delta-1 proteins	
CC	can be used to treat or prevent disorders characterised by	
CC	increased Notch activity, such as cervical, breast, lung or colon	
CC	cancer, melanoma or seminoma, and nervous system disorders or to	
CC	promote tissue regeneration and repair.	
SO	Sequence 740 AA:	
Query Match	44.7%: Score 775; DB 21; Length 740;	
Best Local Similarity	49.1%: Pred. No. 2,256-70;	
Matches 104; Conservative 33; Mismatches 74; Indels 1; Gaps 1;		
Db	20 qvdsqgvefkqlqefvnhkqllsnrnccrpggpgagqgqcdcktfvvciklhyqasvsp 79	
Oy	19 QVHSSGSEFLRLKRYNSNDHGRNDEGRCCSGSDGATGKCLASCKTRFVCLKHQYADITD 78	
Db	80 eppectysaieicpvygansfsvpdqaggaadpafsmipiripfiftvpgtfsllieelhdsp 139	

Dy				:	: :			:	
Dz	TSGCTGCVITPILIGENSVINLDAQRFQNKGFETNPIDPFPSFMSNPGFFSLIVEAMH-DTN	137							
Dc		140	dclttemperllsr atcrhlavgeewsqdlhsgrtdlkysr fvcdehyegscvfc	199					
Dd		138	NSGNARNTKKLLIQRLLVQQVLVESEWTKNSESOYTSLEYDFRVTCDLNLYGSCKAKFC	197					
De		200	rprdrfghtcgcegcgvcmppwqgctcp	231					
Df		196	RPRDSFGHSSTCEGTGCILLGMGDYCHIP	229					
Result		5							
ID	Ml8353 standard; protein; 723 AA.								
AC	WL8353:								
DT	11-FEB-1998 (first entry)								
DE	Proliferation and differentiation suppression polypeptide.								
KM	Proliferation; differentiation; suppression; human; delta-1;								
KX	serrate-1; blood cell; neuron; leukaemia; malignant tumour;								
OS	Immunosuppression.								
FH	Homo sapiens.								
FT	Key								
FT	Peptide								
FT	/label-Signal								
FT	Protein								
FN	MO9J19172-A1.								
PD	29-MAY-1997.								
PE	15-NOV-1996; J03356.								
PR	30-NOV-1995; JP-311811.								
PR	17-NOV-1995; JP-299611.								
PA	(ASAH) ASahi KASEI KOgyo KK.								
PI	Itoh A, Sakano S,								
DR	WPt: 97-238110/27.								
N-PDB:	T70114.								
PT	Peptide(s) encoded by human genes delta-1 and serrate-1 - suppress proliferation and differentiation of undifferentiated human blood cells								
ES	Claim 15; Page 77-82; 11app; Japanese.								
CC	The present sequence represents a polypeptide which suppresses proliferation and differentiation of undifferentiated cells such as neurons and blood cells. The polypeptide may be used for the prevention and control of disorders involving undifferentiated cells, such as leukemia and malignant tumors, and improvement of blood formation. e.g. after immunosuppression.								
SQ	Sequence 723 AA:								
Query Match	43.3%; Score 751; DB 25; Length 723;								
Best Local Similarity	46.5% Pred. No. 9,25e-68;								
Matches	105; Conservative 39; Mismatches 76; Indels 6; Gaps 5;								
Db		4	tcaalajalsallqvassgyfelkqlgefynkg1jgnrnccrgga-gppp-c--acrif	59					
Oy		5	kclltaficftviyoqhssgsfeldrlkyfsndGRDNEGCCSGEDSGATGKCIGSKCTR	64					
Dbb		60	frcvlckhhyaasveepoctygsvatvp1ylvdsflpdggv-adafaenplrfgfwtwp	118					
Oy		65	fRVCLKHQAQTIDTSSOCTGVDTTPILIGENSVNLTDAPRNKGNNFLIOFPFSFWG	124					
Dbb		119	tfslieahldspddlatemperlsiarlarqnrlvtgeeagsqdlhsagrtidlksyrfv	178					
Oy		125	TFSLIVIEWMH-DINNGSNARNKLLIORLVQOVLEVSEWKINKKSQYTLSLEDPFRVT	183					
Ddb		179	cdehyegscsvferpridafightcgegcgvcmppwkpyctec	224					
Oy		184	CGLNWYGSCAFRCRPNDSTRGHSTCSETGTICLLTLTMGGDYCHIP	229					
RESULT		6							
ID	Ml1720 standard; Protein; 722 AA.								
AC	ML1720:								
DT	28-APR-1997 (first entry)								
DE	M-delta-1 polypeptide.								

MUSE (TM)

Release 3.1A John F. Collins, Bioinformatics Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPSrch_PP protein - protein database search, using Smith-Waterman algorithm
Run on: Fri Jun 12 12:06:43 1998; Maspar time 2.40 Seconds
Tabular output not generated. 549.166 Million cell updates/sec

Title: >US-08-083-590A-2
Description: (1-230) from US08083590A.pep (2 of 2)
Perfect Score: 1734
Sequence: 1 MHWKICLITFAICTYIVQV.....ETGELICLTGMOGYCHLPK 230

Scoring table: PAM 150
Gap 11

Searched: 62624 segs, 5720858 residues

Post-processing: Minimum Match 08
Listing first 45 summaries

Database: a-issued
1:5_COMB 2:PCT9_COMB 3:backfiles1

Statistics: Mean 30.190; Variance 125.451; scale 0.241

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	1734	100.0	833	1	US-08-264-Sequence 6, Applicatio	8.30e+164
2	1514	87.3	203	1	US-08-264-Sequence 3, Applicatio	3.50e+141
3	480	27.7	293	1	US-08-264-Sequence 9, Applicatio	3.16e+36
4	452	26.1	199	1	US-08-264-Sequence 4, Applicatio	1.84e+33
5	100	5.8	383	1	US-08-457-Sequence 2, Applicatio	1.11e+00
6	100	5.8	383	1	US-08-597-Sequence 2, Applicatio	1.11e+00
7	95	5.5	385	1	US-08-597-Sequence 1, Applicatio	2.73e+00
8	95	5.5	385	1	US-08-457-Sequence 1, Applicatio	2.73e+00
9	94	5.4	969	2	PCT-US93-0-Sequence 2, Applicatio	3.26e+00
10	92	5.3	768	1	US-08-454-Sequence 4, Applicatio	4.65e+00
11	92	5.3	769	1	US-08-454-Sequence 6, Applicatio	4.65e+00
12	88	5.1	316	1	US-08-260-Sequence 18, Applicati	9.36e+00
13	88	5.1	316	2	PCT-US94-0-Sequence 18, Applicati	9.36e+00
14	88	5.1	316	1	US-08-017-Sequence 4, Applicatio	1.11e+01
15	87	5.0	271	1	US-08-447-Sequence 4, Applicatio	1.11e+01
16	87	5.0	718	1	US-08-444-Sequence 4, Applicatio	1.11e+01
17	87	5.0	718	1	US-08-445-Sequence 4, Applicatio	1.11e+01
18	86	5.0	1810	2	PCT-US95-1-Sequence 4, Applicatio	1.32e+01
19	83	4.8	1124	2	PCT-US93-0-Sequence 2, Applicatio	2.21e+01
20	83	4.8	1124	1	US-08-323-Sequence 2, Applicatio	2.21e+01
21	83	4.8	1833	2	PCT-US95-0-Sequence 18, Applicatio	2.21e+01
22	83	4.8	2199	2	PCT-US95-1-Sequence 2, Applicatio	2.21e+01
23	82	4.7	283	1	US-08-658-Sequence 2, Applicatio	2.61e+01

RESULT	ID	US-08-264-534-6	STANDARD:	PRT:	833 AA.
24	81	4.7	416	1	US-08-252-Sequence 2, Applicatio
25	81	4.7	464	1	US-08-252-Sequence 6, Applicatio
26	81	4.7	925	1	US-08-252-Sequence 4, Applicatio
27	81	4.7	1358	1	US-08-404-Sequence 4, Applicatio
28	81	4.7	1358	1	US-08-404-Sequence 4, Applicatio
29	81	4.7	1358	1	US-08-404-Sequence 4, Applicatio
30	79	4.6	650	1	US-08-325-Sequence 56, Applicati
31	79	4.6	670	1	US-08-477-Sequence 3, Applicatio
32	79	4.6	670	1	US-08-243-Sequence 3, Applicatio
33	79	4.6	670	1	US-08-484-Sequence 3, Applicatio
34	79	4.6	688	1	US-08-325-Sequence 57, Applicati
35	79	4.6	769	1	US-08-484-Sequence 4, Applicatio
36	79	4.6	769	1	US-08-243-Sequence 4, Applicatio
37	79	4.6	769	1	US-08-477-Sequence 4, Applicatio
38	78	4.5	501	1	US-08-434-Sequence 2, Applicatio
39	78	4.5	501	1	US-07-687-Sequence 2, Applicatio
40	78	4.5	501	1	US-08-271-Sequence 2, Applicatio
41	78	4.5	501	1	PCT-US95-0-Sequence 2, Applicatio
42	78	4.5	541	2	US-08-377-Sequence 30, Applicati
43	77	4.4	336	1	US-08-093-Sequence 2, Applicatio
44	77	4.4	336	1	US-07-704-Sequence 3, Applicatio
45	77	4.4	336	1	US-07-704-Sequence 3, Applicatio

ALIGNMENTS

Sequence 6, Application US/08264534.
Patent No. 5648464
GENERAL INFORMATION:
APPLICANT: Artavanis-Tsakonas, Spyridon et al.
TITLE OF INVENTION: Human No. 5648464ch And Delta, Binding Domains
TITLE OF INVENTION: In Topolythmic Proteins, And Methods Based Thereon
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennine & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/264,534
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/695,189
FILING DATE: 03-MAY-1991
ATTORNEY/AGENT INFORMATION:
NAME: MISTOCK, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7326-004
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 790-9090
TELEFAX: 212 8698864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 833 amino acids
TYPE: amino acid

CC TOPOLOGY: unknown
CC MOLECULE TYPE: protein
SQ SEQUENCE 833 AA; 88812 MW; 3155492 CN;

Query Match 100.0%; Score 1734; DB 1; Length 833;
Best Local Similarity 100.0%; Pred. No. 8.30e-164;
Matches 230; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 MHWKCLTARICTVIQVHSSSFEURLXYFSDHGRDNEGRCCSGESDGTGKCLGS 60
QY 1 MHWKCLTARICTVIQVHSSSFEURLXYFSDHGRDNEGRCCSGESDGTGKCLGS 60

Db 61 CKTRFVCLKHQATIDTTSOCTGVDVITPILGENSVNLTDQRFQNGFTNPIQFPFSF 120
QY 61 CKTRFVCLKHQATIDTTSOCTGVDVITPILGENSVNLTDQRFQNGFTNPIQFPFSF 120

Db 121 SWPFTSLIVEAWHDTNNSGNARTNKLIIQRLVQVLEVSSEWTKNKSQYTSLEYDF 180
QY 121 SWPFTSLIVEAWHDTNNSGNARTNKLIIQRLVQVLEVSSEWTKNKSQYTSLEYDF 180

Db 181 RVTCDLNYGSGCAKFCRPRDSEFGHSTCSETGEIICLTGWQGDYCHIPK 230
QY 181 RVTCDLNYGSGCAKFCRPRDSEFGHSTCSETGEIICLTGWQGDYCHIPK 230

RESULT 2
ID US-08-264-534-3 STANDARD; PRT; 203 AA.
XX
AC xxxxxx
XX
DT 01-JAN-1900
XX
DE Sequence 3, Application US/08264534.
XX
CC Sequence 3, Application US/08264534
CC Patent No. 5648464
CC GENERAL INFORMATION:
CC APPLICANT: Artavanis-Tsakonas, Spyridon et al.
CC TITLE OF INVENTION: Human No. 5648464ch And Delta, Binding Domains
CC TITLE OF INVENTION: In Toporythmic Proteins, And Methods Based Thereon
CC NUMBER OF SEQUENCES: 34
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Pennie & Edmonds
CC STREET: 1155 Avenue of the Americas
CC CITY: New York
CC STATE: New York
CC COUNTRY: U.S.A.
CC ZIP: 10036
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/264,534
CC FILING DATE:
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/695,189
CC FILING DATE: 03-MAY-1991
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Misrock, S. Leslie
CC REGISTRATION NUMBER: 18,872
CC REFERENCE/DOCKET NUMBER: 7326-004
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 212 790-9090
CC TELEFAX: 212 8698864/9741
CC TELEX: 66141 PENNIE
CC INFORMATION FOR SEQ ID NO: 3:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 203 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: single

CC TOPOLOGY: unknown
CC MOLECULE TYPE: peptide
SQ SEQUENCE 203 AA; 22840 MW; 218129 CN;

Query Match 87.3%; Score 1514; DB 1; Length 203;
Best Local Similarity 100.0%; Pred. No. 3.50e-141;
Matches 203; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 GSFEURLXYFSDHGRDNEGRCCSGESDGTGKCLGSKTRFVCLKHQATIDTTSOCT 60
QY 24 GSFEURLXYFSDHGRDNEGRCCSGESDGTGKCLGSKTRFVCLKHQATIDTTSOCT 83

Db 61 YGDVITPILGENSVNLTDQRFQNGFTNPIQFPFSFSGPFTSLIVEAWHDTNNSGNAR 120
QY 84 YGDVITPILGENSVNLTDQRFQNGFTNPIQFPFSFSGPFTSLIVEAWHDTNNSGNAR 143

Db 121 TNKLLIQRLLVQVLEVSSEWTKNKSQYTSLEYDFRVTCDLNYGSGCAKFCRPRD 180
QY 144 TNKLLIQRLLVQVLEVSSEWTKNKSQYTSLEYDFRVTCDLNYGSGCAKFCRPRD 203

Db 181 FGHSTCSETGEIICLTGWQGDY 203
QY 204 FGHSTCSETGEIICLTGWQGDY 226

RESULT 3
ID US-08-264-534-9 STANDARD; PRT; 293 AA.
XX
AC xxxxxx
XX
DT 01-JAN-1900
XX
DE Sequence 9, Application US/08264534.
XX
CC Sequence 9, Application US/08264534
CC Patent No. 5648464
CC GENERAL INFORMATION:
CC APPLICANT: Artavanis-Tsakonas, Spyridon et al.
CC TITLE OF INVENTION: Human No. 5648464ch And Delta, Binding Domains
CC TITLE OF INVENTION: In Toporythmic Proteins, And Methods Based Thereon
CC NUMBER OF SEQUENCES: 34
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Pennie & Edmonds
CC STREET: 1155 Avenue of the Americas
CC CITY: New York
CC STATE: New York
CC COUNTRY: U.S.A.
CC ZIP: 10036
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/264,534
CC FILING DATE:
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/695,189
CC FILING DATE: 03-MAY-1991
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Misrock, S. Leslie
CC REGISTRATION NUMBER: 18,872
CC REFERENCE/DOCKET NUMBER: 7326-004
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 212 790-9090
CC TELEFAX: 212 8698864/9741
CC TELEX: 66141 PENNIE
CC INFORMATION FOR SEQ ID NO: 9:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 293 amino acids
CC TYPE: amino acid
CC TOPOLOGY: unknown

cc TOPLOGY: unknown

CC	ORIGINAL SOURCE:
CC	STRAIN: Human Dik
SO	SEQUENCE 383 AA: 41201 MW: 758273 GN

cc TOPLOGY: unknown

CC	ORIGINAL SOURCE:
CC	STRAIN: Human Dik
SO	SEQUENCE 383 AA: 41201 MW: 758273 GN

```

Query Match          5.8%; Score 100; DB 1; Length 383;
Best Local Similarity 32.6%; Pred.No.1.lle+00;
Matches 14; Conservative 8; Mismatches 19; Indels 2; Gaps 2;

Db      45 CQPGWQGPLCDQ-CVTPSGCL-HGLCGEPGQCICTDGWDGELC 85
        ||| : | | : | | : | | : | | : | | : | | : | |
QY      184 CDLNYGSGCAKFCRPDDSFGHSTCSETGEIICLTGWQGDYC 226

RESULT      6
ID US-08-597-545-2 STANDARD; PRT; 383 AA.
XX
XX xxxxxx
XX
XX 01-JAN-1900
XX
DE Sequence 2, Application US/08597545.
XX
CC Sequence 2, Application US/08597545
CC Patent No. 5580738
CC GENERAL INFORMATION:
CC APPLICANT: LABORDA, Jorge
CC TITLE OF INVENTION: Delta-Like Gene Expressed In
CC TITLE OF INVENTION: Neuroendocrine Tumors
CC NUMBER OF SEQUENCES: 9
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Foley & Lardner
CC STREET: 3000 K Street, N.W., Suite 500
CC CITY: Washington, D.C.
CC COUNTRY: USA
CC ZIP: 20007-5109
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/597,545
CC FILING DATE:
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/989,537
CC FILING DATE: 11-DEC-1992
CC ATTORNEY/AGENT INFORMATION:
CC NAME: BENT, Stephen A.
CC REGISTRATION NUMBER: 29,768
CC REFERENCE/DOCKET NUMBER: 40399/166 NIHD
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (202)672-5300
CC TELEFAX: (202)672-5399
CC TELEX: 904136
CC INFORMATION FOR SEQ ID NO: 2:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 383 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC ORIGINAL SOURCE:
CC STRAIN: Human Dlk
CC
SQ SEQUENCE 383 AA; 41201 MW; 758273 CN;

Query Match          5.8%; Score 100; DB 1; Length 383;
Best Local Similarity 32.6%; Pred.No.1.lle+00;
Matches 14; Conservative 8; Mismatches 19; Indels 2; Gaps 2;

Db      45 CQPGWQGPLCDQ-CVTPSGCL-HGLCGEPGQCICTDGWDGELC 85
        ||| : | | : | | : | | : | | : | | : | | : | |
QY      184 CDLNYGSGCAKFCRPDDSFGHSTCSETGEIICLTGWQGDYC 226

RESULT      7
ID US-08-597-545-1 STANDARD; PRT; 385 AA.
XX

```

CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Foley & Lardner
CC STREET: 3000 K Street, N.W., Suite 500
CC CITY: Washington, D.C.
CC COUNTRY: USA
CC ZIP: 20007-5109
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC OPERATING SYSTEM: IBM PC compatible
CC SOFTWARE: Patent Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/457,135
CC FILING DATE: 01-JUN-1995
CC PRIORITY APPLICATION DATA:
CC APPLICATION NUMBER: 07/989,537
CC FILING DATE: 11-DEC-1992
CC ATTORNEY/AGENT INFORMATION:
CC NAME: BENT, Stephen A.
CC REGISTRATION NUMBER: 29,768
CC REFERENCE/DOCKET NUMBER: 40399/304/NIND
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (202)672-5300
CC TELEFAX: (202)672-5399
CC TELEX: 904136
CC INFORMATION FOR SEQ ID NO: 1:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 385 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC ORIGINAL SOURCE:
CC STRAIN: Mouse Dlx
CC SEQ SEQUENCE 385 AA; 41320 MW; 772804 CN;

Query Match 5.5%; Score 95; DB 1; Length 385;
Best Local Similarity 30.5%; Pred. No. 2,73e+00;
Matches 18; Conservative 9; Mismatches 30; Indels 2; Gaps 2;

Db 31 DPQYGFCEADNVCRCYHGWEGPLCDK-CYTAFCV-NGVCKEPMQICIKGMDGKFCFI 87
QY 170 ESQYTSLEDFRVTCDLNYYGSGCAKFCRPDSSFHSTCSETEICLTGMQGDYCHI 228

RESULT 9
ID PCT-US93-02147A-2 STANDARD; PRT; 969 AA.
XX
AC xxxxxx
XX
DT 01-JAN-1900
XX
DE Sequence 2, Application PC/TUS9302147A.
XX
CC Sequence 2, Application PC/TUS9302147A.
CC GENERAL INFORMATION:
CC APPLICANT: BARR, PHILIP J
CC APPLICANT: KEEFER, MICHAEL C
CC TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR PACE 4 AND
CC TITLE OF INVENTION: PACE 4.1 GENE AND POLYPEPTIDES IN CELLS
CC NUMBER OF SEQUENCES: 16
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: COOLEY GODWARD CASTRO HUDDLESON & TATUM
CC STREET: FIVE PALO ALTO SQUARE
CC CITY: PALO ALTO
CC STATE: CALIFORNIA
CC COUNTRY: USA
CC ZIP: 94306
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC OPERATING SYSTEM: IBM PC compatible
CC SOFTWARE: Patent Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US93/02147A

CC FILING DATE: 19930309
CC CLASSIFICATION:
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/848,629
CC FILING DATE: 09-MAR-1992
CC ATTORNEY/AGENT INFORMATION:
CC NAME: NEBLEY PH.D., RICHARD L
CC REGISTRATION NUMBER: 30092
CC REFERENCE/DOCKET NUMBER: CHIR-009/00US
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 494-7622
CC TELEFAX: (415) 857-0663
CC TELEX: 380816 COOLEY PA
CC INFORMATION FOR SEQ ID NO: 2:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 969 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQ SEQUENCE 969 AA; 106419 MW; 4660240 CN;

Query Match 5.4%; Score 94; DB 2; Length 969;
Best Local Similarity 38.6%; Pred. No. 3.26e+00;
Matches 17; Conservative 6; Mismatches 19; Indels 2; Gaps 2;

Db 733 CPLGFEDTARRCR-RCHR-GCETCSSRAATQCLSGRGVYHH 774
QY 184 CDLNYGSGCAKFCRPDSSFHSTCSETEICLTGMQGDYCH 227

RESULT 10
ID US-08-454-455-4 STANDARD; PRT; 768 AA.
XX
AC xxxxxx
XX
DT 01-JAN-1900
XX
DE Sequence 4, Application US/08454455.
XX
CC Sequence 4, Application US/08454455
CC Patent No. 5635601
CC GENERAL INFORMATION:
CC APPLICANT: Moyle, Mathew
CC APPLICANT: McLean, John W.
CC TITLE OF INVENTION: NOVEL BETA INTEGRIN SUBUNIT
CC NUMBER OF SEQUENCES: 9
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Genentech, Inc.
CC STREET: 460 Point San Bruno Blvd
CC CITY: South San Francisco
CC STATE: California
CC COUNTRY: USA
CC ZIP: 94080
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: 3.5 inch, 720 Kb floppy disk
CC OPERATING SYSTEM: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: WinPatIn (Genentech)
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/454,455
CC FILING DATE: 30-MAY-1995
CC CLASSIFICATION: 530
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 08/193989
CC FILING DATE: 09-FEB-1994
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 08/004142
CC FILING DATE: 13-JAN-1993
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 07/670607
CC FILING DATE: 14-MAR-1991
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Lee, Wendy M.

OY 112 NP1QPFPSFWPGTFSLIVEAMHDITNNSGNARTNKL 148

RESULT 13
ID PCT-US94-02034-18 STANDARD; PRT; 316 AA.
XX xxxxxx
AC
DT 01-JAN-1900
DE Sequence 18, Application PC/TUS9402034.
XX Sequence 18, Application PC/TUS9402034
CC GENERAL INFORMATION:
CC APPLICANT: STREET: 201 West 7th Street
CC APPLICANT: CITY: Austin
CC APPLICANT: STATE: Texas
CC APPLICANT: COUNTRY: United States of America
CC APPLICANT: POSTAL CODE: 78701
CC APPLICANT: TELEPHONE NO: (512)499-4462
CC APPLICANT: TELEFAX: (512)499-4523
CC TITLE OF INVENTION: DETECTION OF CONVERSION TO MUCCOYD IN PSEUDOMONAS
CC TITLE OF INVENTION: AERUGINOSA INFECTING CYSTIC FIBROSIS PATIENTS
CC NUMBER OF SEQUENCES: 23
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Arnold, White & Durkee
CC STREET: P. O. Box 4433
CC CITY: Houston
CC STATE: Texas
CC COUNTRY: USA
CC ZIP: 77210
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: FLOPPY DISK
CC COMPUTER: IBM PC COMPATIBLE
CC OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US94/02034
CC FILING DATE: UNKNOWN
CC CLASSIFICATION:
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 08/017,114
CC FILING DATE: 12 February 1993 (12.02.93)
CC ATTORNEY/AGENT INFORMATION:
CC NAME: HODGINS, DANIEL S.
CC REGISTRATION NUMBER: 31,026
CC REFERENCE/DOCKET NUMBER: UTPK231P--
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 512/320-7200
CC TELEFAX: 713/789-2679
CC TELEX: 79-0924
CC INFORMATION FOR SEQ ID NO: 18:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 316 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: peptide
CC SEQUENCE 316 AA; 34572 MW; 501312 CN;
Query Match 5.1%; Score 88; DB 2; Length 316;
Best Local Similarity 35.1%; Pred. No. 9.36e+00;
Matches 13; Conservative 9; Mismatches 13; Indels 2; Gaps 2;
DB 38 NSFQGTVEYERNGSFS-THEIWHRVESDGAIVR-ERLL 72
OY 112 NP1QPFPSFWPGTFSLIVEAMHDITNNSGNARTNKL 148
RESULT 14
ID US-08-017-114-18 STANDARD; PRT; 316 AA.
XX

AC xxxxxx
XX
DT 01-JAN-1900
DE Sequence 18, Application US/08017114.
XX Sequence 18, Application US/08017114
CC Patent No. 5591838
CC GENERAL INFORMATION:
CC APPLICANT: Dericic, Vojko
CC APPLICANT: Martin, Daniel W.
CC TITLE OF INVENTION: DETECTION OF CONVERSION TO MUCCOYD IN
CC TITLE OF INVENTION: PSEUDOMONAS AERUGINOSA INFECTING CYSTIC FIBROSIS PATIEN
CC NUMBER OF SEQUENCES: 19
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Arnold, White & Durkee
CC STREET: P. O. Box 4433
CC CITY: Houston
CC STATE: Texas
CC COUNTRY: USA
CC ZIP: 77210
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent In Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/017,114
CC FILING DATE: 19930212
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Hodgins, Daniel S.
CC REGISTRATION NUMBER: 31,026
CC REFERENCE/DOCKET NUMBER: UTSK:205/HOD
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 512/320-7200
CC TELEFAX: 512/474-7577
CC TELEX: N/A
CC INFORMATION FOR SEQ ID NO: 18:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 316 amino acids
CC TYPE: AMINO ACID
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: peptide
CC SEQUENCE 316 AA; 34572 MW; 501312 CN;
Query Match 5.1%; Score 88; DB 1; Length 316;
Best Local Similarity 35.1%; Pred. No. 9.36e+00;
Matches 13; Conservative 9; Mismatches 13; Indels 2; Gaps 2;
DB 38 NSFQGTVEYERNGSFS-THEIWHRVESDGAIVR-ERLL 72
OY 112 NP1QPFPSFWPGTFSLIVEAMHDITNNSGNARTNKL 148
RESULT 15
ID US-08-447-554-4 STANDARD; PRT; 271 AA.
XX
AC xxxxxx
DT 01-JAN-1900
DE Sequence 4, Application US/08447554.
XX Sequence 4, Application US/08447554
CC Patent No. 5661003
CC GENERAL INFORMATION:
CC APPLICANT: FUSHIMI, KIYOHIDE
CC APPLICANT: UCHIDA, SHINICHI
CC APPLICANT: SASAKI, SEI
CC APPLICANT: MARUMO, FUMIAKI
CC TITLE OF INVENTION: WATER CHANNEL

Release 3.1a John F. Collins, Bioinformatics Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

(TM)

Release 3.1a John F. Collins, Bioinformatics Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MSrch_pp protein - protein database search, using Smith-Waterman algorithm
Run on: Fri Jun 12 12:03:56 1998; Maspar time 10.84 Seconds
Tabular output not generated.

Title: >US-08-083-590A-2
Description: (1-230) from US08083590A.pep (2 of 2)
Perfect Score: 1734
Sequence: 1 MHWIKCLTAFICFTVIYQV.....ETGEICLTGMOGDYCHIRP 230

Scoring table:
PAM 150
Gap 11

Searched: 120441 seqs, 36531193 residues

Post-processing: Minimum Match 08
Listing first 45 summaries

Database: p1r56
1:p1r1 2:p1r2 3:p1r3 4:p1r4 5:mr13d

Statistics: Mean 42.868; Variance 79.045; scale 0.542

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed.
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	1734	100.0	833	2	gene Delta protein pr	0.00e+00
2	1728	99.7	832	2	neurogenic protein De	0.00e+00
3	1722	99.3	880	2	gene Delta protein pr	0.00e+00
4	775	44.7	728	2	C-Delta-1 - chicken	1.35e-141
5	732	42.2	722	2	DELTA-like 1 - mouse	1.35e-131
6	568	32.8	1220	2	Jagged protein precu	2.06e-95
7	480	27.7	1404	2	serate protein precu	2.63e-76
8	480	27.7	1408	2	gene serrate protein	2.63e-76
9	133	7.7	402	2	lag-2 protein - Caeno	7.32e-07
10	104	6.0	383	2	delta-like dlk homeo	3.29e-02
11	103	5.9	788	2	Integrin beta-3 subu	8.83e-02
12	100	5.8	259	2	fetal antigen 1 - hum	8.83e-02
13	100	5.8	260	2	fetal antigen 1 home	8.83e-02
14	100	5.8	383	2	homeotic protein dlk	8.83e-02
15	101	5.8	788	2	Integrin beta-6 chain	6.37e-02
16	101	5.8	2471	2	cell-fate determining	2.33e-01
17	97	5.6	301	2	hypothetical protein	2.33e-01
18	97	5.6	377	2	Integrin beta-6 chain	2.33e-01
19	95	5.5	385	2	homeotic protein dlk	4.39e-01
20	96	5.5	781	2	Integrin beta3 - chic	3.20e-01
21	96	5.5	3566	2	tenascin-X precursor	3.20e-01
22	94	5.4	810	2	f42h10.5 protein - Ca	6.01e-01
23	94	5.4	962	2	subtilisin-like prote	6.01e-01

24	94	5.4	969	2	A39490	serine proteinase (EC	6.01e-01
25	94	5.4	975	2	JC5570	subtilisin-like prote	6.01e-01
26	94	5.4	1203	2	A49175	Morch B protein - mou	6.01e-01
27	92	5.3	590	1	QRAYPR	arginine transport pr	1.12e+00
28	92	5.3	768	2	B41029	Integrin beta-8 chain	1.12e+00
29	92	5.3	769	2	A41029	Integrin beta-8 chain	1.12e+00
30	90	5.2	522	2	JT0980	cytochrome P450 52A2	2.06e+00
31	91	5.2	858	5	10BA	chitinase n-acetyl-b	1.52e+00
32	91	5.2	858	5	10BB	chitinase (EC 3.2.1.	1.52e+00
33	91	5.2	885	3	JC4732	beta-N-Acetylhexosam	1.52e+00
34	90	5.2	937	2	I53282	gene PAC4 protein -	2.06e+00
35	90	5.2	1820	2	A55494	latent transforming g	3.77e+00
36	88	5.1	47	2	S11530	purothionin gamma -	3.77e+00
37	88	5.1	316	2	A47064	align protein - Pseudo	2.79e+00
38	89	5.1	408	2	B64708	glutamate permease -	3.77e+00
39	88	5.1	715	2	S76492	hypothetical protein	3.77e+00
40	88	5.1	782	2	A61625	tenascin-like protein	3.77e+00
41	87	5.0	92	2	D37057	epithelial cell glyco	5.07e+00
42	87	5.0	271	2	JT0750	water channel protein	5.07e+00
43	87	5.0	788	2	A26547	platelet glycoprotein	5.07e+00
44	87	5.0	1928	2	J50610	beta-galactosidase (E	5.07e+00
45	87	5.0	2515	2	S47008	tenascin-like protein	5.07e+00

ALIGNMENTS

RESULT 1
ENTRY S19087 #type complete
TITLE gene Delta protein precursor - fruit fly (Drosophila melanogaster)
ORGANISM #formal_name Drosophila melanogaster
DATE 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 16-Feb-1997

ACCESSIONS S19087
REFERENCE S19087
#authors Muskavitch, M.A.T.
#submission submitted to the EMBL Data Library, June 1991
#accession S19087
#status preliminary
#molecule_type mRNA
#residues 1-833 #label MUS
#cross-references EMBL:Y00222

GENETICS
#gene FlyBase:D1
#cross-references FlyBase:FBgn0000463

SUMMARY
#length 833 #molecular_weight 88812 #checksum 2566

Query Match	100.0%	Score 1734	DB 2	Length 833
Best Local Similarity 100.0%	Pred. No. 0.00e+00	Mismatches 0	Indels 0	Gaps 0
Matches 230	Conservative 0	Mismatches 0	Indels 0	Gaps 0
Db	1	MHWIKCLTAFICFTVIYQVHSSGFELRLKFSNDGRNDRGRCSSGSDGATGCLGS	60	
QY	1	MHWIKCLTAFICFTVIYQVHSSGFELRLKFSNDGRNDRGRCSSGSDGATGCLGS	60	
Db	61	CKTRFVCLKHYQATIDTTSQCTGVDVITPILGENSVNLDAQRQNGKFTNPLOPPSF	120	
QY	61	CKTRFVCLKHYQATIDTTSQCTGVDVITPILGENSVNLDAQRQNGKFTNPLOPPSF	120	
Db	121	SWPGTFSILVAMHDNNSGNAKRTKLLIQLVQVLEVSSEWTKNSSEQYTSLEYDF	180	
QY	121	SWPGTFSILVAMHDNNSGNAKRTKLLIQLVQVLEVSSEWTKNSSEQYTSLEYDF	180	
Db	181	RVTCDLNYGSGCAKFCRPRDSFGHSCTSGEITCLTGMGDYCHIRP 230		
QY	181	RVTCDLNYGSGCAKFCRPRDSFGHSCTSGEITCLTGMGDYCHIRP 230		
RESULT 2	A31246	#type complete		
ENTRY	neurogenic protein Delta precursor - fruit fly (Drosophila melanogaster)			
TITLE	#formal_name Drosophila melanogaster			
ORGANISM				

```
DATE 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change
16-Feb-1997
ACCESSIONS A31246
REFERENCE A31246
#authors Kopynski, C.C.; Alton, A.K.; Fichtel, K.; Kooh, P.J.;
#journal Muskavitch, M.A.T.
#title Genes Dev. (1988) 2:1723-1735
#cross-references MUID:89196890
#accession A31246
#molecule_type mRNA
#residues_type 1-832 #label KOP
#cross-references GB:Y00222
GENETICS
#gene FlyBase:D1
#cross-references FlyBase:FBgn0000463
SUMMARY
#length 832 #molecular-weight 88943 #checksum 636
Query Match 99.7%; Score 1728; DB 2; Length 832;
Best Local Similarity 99.6%; Pred. No. 0.00e+00;
Matches 229; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Db 1 MHWIKLLTAFICFTVIVQVHSSGSFELRLKYFSNDHRDNEGRCCSGESDGTGKCLGS 60
|||||
Qy 1 MHWIKLLTAFICFTVIVQVHSSGSFELRLKYFSNDHRDNEGRCCSGESDGTGKCLGS 60
|||||
Db 61 CKTRFRLCLKHYQATIDTTTSCQTYGDIPTILGNSVNLTDQRFQNGKFTNPQPPFSF 120
|||||
Qy 61 CKTRFRLCLKHYQATIDTTTSCQTYGDIPTILGNSVNLTDQRFQNGKFTNPQPPFSF 120
|||||
Db 121 SWPGTFSLIVEAWHDNNSGNARTKLLIQRLLVQVLEVSSEWKNKSSQTSLEYDF 180
|||||
Qy 121 SWPGTFSLIVEAWHDNNSGNARTKLLIQRLLVQVLEVSSEWKNKSSQTSLEYDF 180
|||||
Db 181 RVTCDLNYGSCAKFCRPRDSDFGHSTCSETGEIICLTGWQGDYCHIPK 230
|||||
Qy 181 RVTCDLNYGSCAKFCRPRDSDFGHSTCSETGEIICLTGWQGDYCHIPK 230
|||||
GENETICS
#gene FlyBase:D1
#cross-references FlyBase:FBgn0000463
SUMMARY
#length 832 #molecular-weight 88943 #checksum 636
Query Match 99.7%; Score 1728; DB 2; Length 832;
Best Local Similarity 99.6%; Pred. No. 0.00e+00;
Matches 229; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Db 1 MHWIKLLTAFICFTVIVQVHSSGSFELRLKYFSNDHRDNEGRCCSGESDGTGKCLGS 60
|||||
Qy 1 MHWIKLLTAFICFTVIVQVHSSGSFELRLKYFSNDHRDNEGRCCSGESDGTGKCLGS 60
|||||
Db 61 CKTRFRLCLKHYQATIDTTTSCQTYGDIPTILGNSVNLTDQRFQNGKFTNPQPPFSF 120
|||||
Qy 61 CKTRFRLCLKHYQATIDTTTSCQTYGDIPTILGNSVNLTDQRFQNGKFTNPQPPFSF 120
|||||
Db 121 SWPGTFSLIVEAWHDNNSGNARTKLLIQRLLVQVLEVSSEWKNKSSQTSLEYDF 180
|||||
Qy 121 SWPGTFSLIVEAWHDNNSGNARTKLLIQRLLVQVLEVSSEWKNKSSQTSLEYDF 180
|||||
Db 181 RVTCDLNYGSCAKFCRPRDSDFGHSTCSETGEIICLTGWQGDYCHIPK 230
|||||
Qy 181 RVTCDLNYGSCAKFCRPRDSDFGHSTCSETGEIICLTGWQGDYCHIPK 230
|||||
RESULTS 3
ENTRY S00670 #type complete
TITLE gene Delta protein precursor - fruit fly (Drosophila melanogaster)
ALTERNATE_NAMES gene D1 protein
ORGANISM #formal_name Drosophila melanogaster
DATE 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change
16-Feb-1997
ACCESSIONS S00670
REFERENCE S00670
#authors Vaessin, H.; Bremer, K.A.; Knust, E.; Campos-Ortega, J.A.
#journal EMBO J. (1987) 6:3431-3440
#title The neurogenic gene Delta of Drosophila melanogaster is expressed in neurogenic territories and encodes a putative transmembrane protein with EGF-like repeats.
#accession S00670
#molecule_type mRNA
#residues 1-880 #label VAE
#cross-references EMBL:X06289
GENETICS
#gene Delta
#cross-references FlyBase:FBgn0000463
KEYWORDS transmembrane protein
FEATURE
1-18 #domain signal sequence #status predicted #label SIG\
19-88 #product gene Delta protein #status predicted #label MAT
SUMMARY #length 880 #molecular-weight 94643 #checksum 7597
Query Match 99.3%; Score 1722; DB 2; Length 880;
Best Local Similarity 99.1%; Pred. No. 0.00e+00;
Matches 226; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
Db 1 MHWIKLLTAFICFTVIVQVHSSGSFELRLKYFSNDHRDNEGRCCSGESDGTGKCLGS 60
|||||
Qy 1 MHWIKLLTAFICFTVIVQVHSSGSFELRLKYFSNDHRDNEGRCCSGESDGTGKCLGS 60
|||||
Db 61 CKTRFRLCLKHYQATIDTTTSCQTYGDIPTILGNSVNLTDQRFQNGKFTNPQPPFSF 120
|||||
Qy 61 CKTRFRLCLKHYQATIDTTTSCQTYGDIPTILGNSVNLTDQRFQNGKFTNPQPPFSF 120
|||||
Db 121 SWPGTFSLIVEAWHDNNSGNARTKLLIQRLLVQVLEVSSEWKNKSSQTSLEYDF 180
|||||
Qy 121 SWPGTFSLIVEAWHDNNSGNARTKLLIQRLLVQVLEVSSEWKNKSSQTSLEYDF 180
|||||
Db 181 RVTCDLNYGSCAKFCRPRDSDFGHSTCSETGEIICLTGWQGDYCHIPK 230
|||||
Qy 181 RVTCDLNYGSCAKFCRPRDSDFGHSTCSETGEIICLTGWQGDYCHIPK 230
|||||
RESULTS 5
ENTRY I48324 #type complete
TITLE DELTA-like 1 - mouse
ORGANISM #formal_name Mus musculus #common_name house mouse
DATE 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change
28-Feb-1997
ACCESSIONS I48324
REFERENCE I48324
#authors Bettenhausen, B.; de Angelis, M.H.; Simon, D.; Guenet, J.L.; Gossler, A.
#journal Development (1995) 121:2407-2418
#title Transient and restricted expression during mouse embryogenesis of Dll1, a murine gene closely related to Drosophila Delta.
```


[illegible]

```

RESULT   12
ENTRY    S48713      #type complete
TITLE    fetal antigen I - human
ORGANISM Homo sapiens #common_name man
DATE     01-Aug-1995 #sequence_revision 10-Nov-1995 #text_change
ACCESSIONS
REFERENCE #authors Jensen, C.H.; Krogh, T.N.; Hojrup, P.; Clausen, P.P.; Skjodt
#journal Eur. J. Biochem. (1994) 225:83-92
#title Protein structure of fetal antigen I (FAI). A novel
circulating human epidermal growth-factor-like protein
expressed in neuroendocrine tumors and its relation to the
gene products of dlk and pgz.
S48713
#accession S48713
#status preliminary
#molecule_type protein
#residues 1-259 ##label JEN
SUMMARY  #length 259 #molecular_weight 27316 #checksum 4381

Query Match          5.8%; Score 100; DB 2; Length 259;
Best Local Similarity 32.6%; Pred. No. 8.83e-02;
Matches 14; Conservative 8; Mismatches 19; Indels 2; Gaps 2;






Db       22 COPGMOGPLCDQ-CVTSRGL-HGLCGEPQCICITDMDIELC 62
        |:|::|||::|::|::|::|::|::|::|::|::|::|::|::|::|
Oy       184 CDLNTYSGCAKFCRPDDSRGFHSTCSFTEIGLICLTGWGDYC 226

RESULT   13
ENTRY    A44549      #type complete
TITLE    fetal antigen I homeotic protein - human
ORGANISM Homo sapiens #common_name man
DATE     03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change
ACCESSIONS
REFERENCE #authors Hojrup, P.; Jensen, C.H.; Skjodt, K.; Teisner, B.
#journal Protein Sci. (1993) 2(Suppl.1):259S
#title Primary structure of human fetal antigen I (FAI), a putative
homeotic glycoprotein.
A44549
#accession A44549
##status preliminary
##molecule_type protein
##residues 1-260 ##label HOJ
SUMMARY  #length 260 #molecular_weight 27413 #checksum 7173

Query Match          5.8%; Score 100; DB 2; Length 260;
Best Local Similarity 32.6%; Pred. No. 8.83e-02;
Matches 14; Conservative 8; Mismatches 19; Indels 2; Gaps 2;

Db       22 COPGMOGPLCDQ-CVTSRGL-HGLCGEPQCICITDMDIELC 62
        |:|::|||::|::|::|::|::|::|::|::|::|::|::|::|::|
Oy       184 CDLNTYSGCAKFCRPDDSRGFHSTCSFTEIGLICLTGWGDYC 226

RESULT   14
ENTRY    S53716      #type complete
TITLE    homeotic protein dlx long form - human
CONTRINS homeotic protein dlx short form
ORGANISM Homo sapiens #common_name man
DATE     06-Dec-1996 #sequence_revision 06-Dec-1996 #text_change
ACCESSIONS
REFERENCE #authors Lee, Y.L.; Helman, L.; Hoffman, T.; Laborda, J.
#journal Biochim. Biophys. Acta (1995) 1261:223-232
#title dlx, pc2 and pre-1 mRNAs encode similar proteins belonging
to the EGF-like superfamily. Identification of polymorphic
variants of this RNA.
S53716
#accession S53716
##molecule_type mRNA
```

[illegible]

FT	CARBOHYD	196	196	POTENTIAL.
FT	CARBOHYD	247	247	POTENTIAL.
FT	CARBOHYD	331	331	POTENTIAL.
FT	CARBOHYD	412	412	POTENTIAL.
FT	CARBOHYD	452	452	POTENTIAL.
FT	CARBOHYD	558	558	POTENTIAL.
FT	CARBOHYD	739	739	POTENTIAL.
FT	CARBOHYD	965	965	POTENTIAL.
FT	CARBOHYD	977	977	POTENTIAL.
FT	CARBOHYD	1004	1004	POTENTIAL.
FT	CARBOHYD	1030	1030	POTENTIAL.
FT	CARBOHYD	1150	1150	POTENTIAL.
FT	CONFLICT	14	17	MISSING (IN REF. 2).
FT	CONFLICT	27	27	P -> A (IN REF. 2).
FT	CONFLICT	1352	1352	T -> S (IN REF. 2).
SEQ	SEQUENCE	1408 AA;	150660 MW; A494A358 CRC32;	

Query Match 27.7%; Score 480; DB 1; Length 1408;
Best Local Similarity 36.1%; Pred.No. 1.25e-93; Indels 16; Gaps 11;
Matches 8; Conservative 59; Mismatches 70;

Db	67	CNLIALI-LILLVHKISAAGNFLEILEISNTNSHLLNGYCCGMPAELRATKTKIGCSPT	125
Qy	6	CLLPAFICFTIVQ-VHSSGSFELRLAYFSNDHGRNDRCCSGESDGAATGKCLG-S-CK	62
Db	126	TAFRLCLKEYOTTTGGASISTGCSFGNATTKILGSSFSVLSDP-----G-VGAIVLPFT	178
Qy	63	TRFRVCLAHYQATID-TT-SQ-CYTGVDVITPILGSESVNLTDQAQFQNKGFNTPIOPFES	119
Db	179	FRWTKSFTLLIQA-LDMYNTSYPDAERLIETSVSGVILP-SPEWTKLDHIGRNATYR	236
Qy	120	FSPGQFTSLIVEAHDNTNSGNARTKLLIQRLLVQQLVLEVSSEWTKNKSEQITSLIED	179
Db	237	VRVQCAVYNTTCTTCFPRDDQFGHYACGSEQKLCGLNGWQVNC	283
Qy	180	FVYTCDLNYGSCAKFCPRDDSFHSTCSETGEILICLTGWQGDYC	226

RESULT	5	STANDARD;	PRT;	402 AA.
ID	LAG2 CAEEL			
AC	P45442;			
DT	01-NOV-1995 (REL. 32, CREATED)			
DT	01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)			
DT	01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)			
DE	LAG-2 PROTEIN PRECURSOR.			
GN	LAG-2 OR LET-461.			
OS	CANORHABDITIS ELEGANS			
OC	EUKARYOTA; METAZOA; ACCELOMATES; NEMATODA; SECERNENTEA; RHABDITIDA.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=BRISTOL N2.			
RX	MEDLINE: 94187845.			
RA	TAX F.E., YEAGERS J.J., THOMAS J.H.;			
RL	NATURE 368:150-154(1994).			
CC	!- FUNCTION: PUTATIVE INTERCELLULAR SIGNAL FOR LIN-12 AND GLP-1 RECEPTORS.			
CC	!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.			
CC	!- SIMILARITY: CONTAINS 1 DSL (DELTA/SERRATE/LAG-2 DOMAIN), WHICH IS REQUIRED TO BIND NOTCH IN DELTA AND SERRATE.			
CC	!- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.			
DR	EMBL; L77495; G459579; -			
DR	PROSITE; PS00022; EGF.1; 2.			
DR	PROSITE; PS01186; EGF.2; 2.			
KW	DIFFERENTIATION; REPEAT; TRANSMEMBRANE; EGF-LIKE DOMAIN; GLYCOPROTEIN; SIGNAL.			
KW	SIGNAL	1	15	POTENTIAL.
FT	CHAIN	16	402	LAG-2 PROTEIN.
FT	DOMAIN	16	279	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	280	306	POTENTIAL.
FT	DOMAIN	307	402	CYTOPLASMIC (POTENTIAL).
FT	DOMAIN	122	166	DSL.
FT	DOMAIN	171	216	EGF-LIKE 1.
FT	DOMAIN	229	266	EGF-LIKE 2.

	Query Match	6.7%	Score 116;	DB 1;	Length 515;
	Best Local Similarity	33.3%;	Pred. No. 1.28e-05;		
	Matches	17;	Conservative	9;	Mismatches 23; Indels 2; Gaps
Dd	127	QLRNCSNSNYGKRCNRKYCIA-NAKL-HWMCSTHGVRCSAGSGEDCSNP	175		
	:				
Oy	179	DFRYTCDLTNGYSGCAFCPRPRODSFHSICSETGELICLTITMOGDYCHIP	229		
RESULT	7	STANDARD:	PRT:	383 AA.	
ID	DLK HUMAN				
AC	P80370;				
DT	01-NOV-1995 (REL. 32, CREATED)				
DI	01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)				
DR	01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)				
DE	DELTA-LIKE PROTEIN PRECURSOR (CONTAINS: FETAL ANTIGEN 1) (FAI) (DLK).				
GN	DLK.				
OS	HOMO SAPIENS (HUMAN).				
OC	EUDAROTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;				
OC	EUTHERIA; PRIMATES.				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=ADRENAL GLAND;				
RX	MEDLINE; 93179372.				
RA	LABORDA J., SAUSVILLE E.A., HOFFMAN T., NOTARIO V.;				
RL	J. BIOL. CHEM. 268:3817-3820(1993).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=PLACENTA, AND ADRENAL GLAND;				
RX	MEDLINE; 95226449.				
RA	LEE Y.L., HELMAN L., HOFFMAN T., LABORDA J.;				
RL	BIOCHIM. BIOPHYS. ACTA 1261:223-232(1995).				
RN	[3]				
RP	SEQUENCE OF 24-383.				
RC	TISSUE=AMNIOTIC FLUID;				
RX	MEDLINE; 95010145.				
RA	JENSEN C.H., KROGH T.N., HOEJTRUP P., CLAUSEN P.P., SKOJD T.K.,				
RL	LARSSON L.-I., ENGHILD J.J., TEISNER B.;				
RN	EUR. J. BIOCHEM. 225:83-92(1994).				
RN	[4]				
RP	SEQUENCE OF 24-60.				
RC	TISSUE=AMNIOTIC FLUID;				
RX	MEDLINE; 93273893.				
RA	JENSEN C.H., TEISNER B., HOJTRUP P., RASMUSSEN H.B., MADSEN O.D.,				
RL	NIELSEN B., SKOJD T.K.;				
RL	HDM. REPROD. 8:635-641(1993).				
CC	-1- FUNCTION: MAY HAVE A ROLE IN NEUROENDOCRINE DIFFERENTIATION.				
CC	-1- SUBUNIT: MONOMER.				
CC	-1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.				
CC	-1- TISSUE SPECIFICITY: FAI IS FOUND WITHIN THE STROMAL CELLS IN CLOSE CONTACT TO THE VASCULAR STRUCTURE OF PLACENTAL VILLI, YOLK SAC, FETAL LIVER, ADRENAL CORTEX AND PANCREAS AND IN THE BETA CELLS OF THE ISLETS OF LANGERHANS IN THE ADULT PANCREAS. FOUND ALSO IN SOME FORMS OF NEUROENDOCRINE LUNG TUMOR TISSUE.				
CC	-1- PTM: N- AND O-GLYCOSYLATED.				
CC	-1- ALTERNATIVE PRODUCTS: AT LEAST ONE SHORTER PROTEIN IS DERIVED FROM THE SAME GENE BY ALTERNATIVE SPLICING.				
CC	-1- SIMILARITY: CONTAINS 6 EGF-LIKE DOMAINS.				
DR	EMBL; Z12172; G38479; -.				
DR	EMBL; U15979; G562106; -.				
DR	EMBL; U15981; G562110; -.				
DR	PROSITE; PS00010; ASX_HYDROXYL, 1.				
DR	PROSITE; PS00022; EGF_1; 5.				
DR	PROSITE; PS0186; EGF_2; 5.				
KW	ANTIGEN; TRANSMEMBRANE; EGF-LIKE DOMAIN; REPEAT; GLYCOPROTEIN; SIGNAL; ALTERNATIVE SPLICING.				
KW	SIGNAL	1	23		
FT	CHAIN	24	383		DELTA-LIKE PROTEIN.
FT	PEPTIDE	24	303		FETAL ANTIGEN 1.
FT	DOMAIN	24	303		EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	304	327		POTENTIAL.
FT	DOMAIN	328	383		CYTOPLASMIC (POTENTIAL).

FT	CARBOHYD	277	277	POTENTIAL.
FT	CARBOHYD	322	322	POTENTIAL.
FT	CARBOHYD	330	330	POTENTIAL.
FT	CARBOHYD	400	400	POTENTIAL.
FT	CARBOHYD	434	434	POTENTIAL.
FT	CARBOHYD	555	555	POTENTIAL.
FT	NON TER	577	577	POTENTIAL.
SO	SEQUENCE	577 AA;	62298 MM;	AE102507 CRC32;
Query Match				
Best Local Similarity		39.5%;	Score 97;	DB 1; Length 577;
Matches		15; Conservative	8; Mismatches	11; Indels 4; Gaps
Db	398 CDNESCVRHKGILCGDNGDC-ECGECCRSRGWGEHCN	434		
Qy	193 CAKE-C-RPRDSEFG-HSTCEGEIICLTGMOGDYCH	227		
RESULT 11				
ID	LM2.CAEEL	STANDARD;	PRT;	3672 AA.
AC	Q21313;			
DT	01-NOV-1997 (REL. 35, CREATED)			
DT	01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)			
DT	01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)			
DE	LAMININ-LIKE PROTEIN K08C7.3 PRECURSOR.			
GN	K08C7.3.			
OS	CAENORHABDITIS ELEGANS.			
OC	EDUAROTA; METAZOA; ACLOLOMATES; NEMATODA; SECERNENTEA; RHABDITIDA			
RN	[1]			
RC	SEQUENCE FROM N.A.			
RP	STRAIN-BRISTOL N2;			
RA	BERKS M.;			
RL	SUBMITTED (MAR-1996) TO EMBL/GENBANK/DBJ DATA BANKS.			
CC	-1- SIMILARITY: CONTAINS ONE LAMININ N-TERMINAL DOMAIN (DOMAIN VI)			
CC	-1- SIMILARITY: CONTAINS 21.5 LAMININ EGF-LIKE DOMAINS.			
CC	-1- SIMILARITY: CONTAINS ONE LAMININ DOMAIN IV.			
DR	EMBL; 470286; E233841; -			
DR	MORPPER; K08C7.3; C06136.			
DR	PROSITE; PS000022; EGF_1; 19.			
DR	PROSITE; PS01186; EGF_2; 4.			
DR	PROSITE; PS01248; LAMININ_TYPE_EGF; 21.			
KW	HYPOTHEETICAL PROTEIN; LAMININ EGF-LIKE DOMAIN; SIGNAL; REPEAT.			
FT	SIGNAL	1	27	POTENTIAL.
FT	CHAIN	28	3672	LAMININ-LIKE PROTEIN K08C7.3.
FT	DOMAIN	28	297	LAMININ N-TERMINAL (DOMAIN VI).
FT	DOMAIN	298	356	LAMININ EGF-LIKE 1.
FT	DOMAIN	357	426	LAMININ EGF-LIKE 2.
FT	DOMAIN	427	471	LAMININ EGF-LIKE 3.
FT	DOMAIN	472	518	LAMININ EGF-LIKE 4.
FT	DOMAIN	519	563	LAMININ EGF-LIKE 5.
FT	DOMAIN	564	609	LAMININ EGF-LIKE 6.
FT	DOMAIN	610	655	LAMININ EGF-LIKE 7.
FT	DOMAIN	656	700	LAMININ EGF-LIKE 8.
FT	DOMAIN	701	755	LAMININ EGF-LIKE 9.
FT	DOMAIN	756	808	LAMININ EGF-LIKE 10.
FT	DOMAIN	809	839	LAMININ EGF-LIKE 11 (INCOMPLETE).
FT	DOMAIN	1415	1460	LAMININ EGF-LIKE 12.
FT	DOMAIN	1461	1505	LAMININ EGF-LIKE 13.
FT	DOMAIN	1506	1553	LAMININ EGF-LIKE 14.
FT	DOMAIN	1554	1604	LAMININ EGF-LIKE 15.
FT	DOMAIN	1605	1614	LAMININ EGF-LIKE 16 (N-TERMINAL).
FT	DOMAIN	1615	1796	LAMININ DOMAIN IV.
FT	DOMAIN	1797	1829	LAMININ EGF-LIKE 16 (C-TERMINAL).
FT	DOMAIN	1830	1879	LAMININ EGF-LIKE 17.
FT	DOMAIN	1880	1936	LAMININ EGF-LIKE 18.
FT	DOMAIN	1937	1989	LAMININ EGF-LIKE 19.
FT	DOMAIN	1990	2036	LAMININ EGF-LIKE 20.
FT	DOMAIN	2037	2083	LAMININ EGF-LIKE 21.
FT	DOMAIN	2084	2131	LAMININ EGF-LIKE 22.
FT	DISULEFD	298	307	BY SIMILARITY.
FT	DISULEFD	300	320	BY SIMILARITY.
FT	DISULEFD	322	331	BY SIMILARITY.
FT	DISULEFD	334	354	BY SIMILARITY.

FT DISULFID 357 366 BY SIMILARITY.
FT DISULFID 359 391 BY SIMILARITY.
FT DISULFID 394 403 BY SIMILARITY.
FT DISULFID 406 424 BY SIMILARITY.
FT DISULFID 427 438 BY SIMILARITY.
FT DISULFID 439 445 BY SIMILARITY.
FT DISULFID 447 456 BY SIMILARITY.
FT DISULFID 459 469 BY SIMILARITY.
FT DISULFID 472 484 BY SIMILARITY.
FT DISULFID 474 491 BY SIMILARITY.
FT DISULFID 493 502 BY SIMILARITY.
FT DISULFID 505 516 BY SIMILARITY.
FT DISULFID 519 531 BY SIMILARITY.
FT DISULFID 521 538 BY SIMILARITY.
FT DISULFID 540 549 BY SIMILARITY.
FT DISULFID 552 561 BY SIMILARITY.
FT DISULFID 564 576 BY SIMILARITY.
FT DISULFID 566 583 BY SIMILARITY.
FT DISULFID 585 594 BY SIMILARITY.
FT DISULFID 597 607 BY SIMILARITY.
FT DISULFID 610 622 BY SIMILARITY.
FT DISULFID 612 629 BY SIMILARITY.
FT DISULFID 631 640 BY SIMILARITY.
FT DISULFID 643 653 BY SIMILARITY.
FT DISULFID 656 668 BY SIMILARITY.
FT DISULFID 658 674 BY SIMILARITY.
FT DISULFID 676 685 BY SIMILARITY.
FT DISULFID 688 698 BY SIMILARITY.
FT DISULFID 701 715 BY SIMILARITY.
FT DISULFID 703 724 BY SIMILARITY.
FT DISULFID 726 735 BY SIMILARITY.
FT DISULFID 728 733 BY SIMILARITY.
FT DISULFID 756 770 BY SIMILARITY.
FT DISULFID 758 777 BY SIMILARITY.
FT DISULFID 779 788 BY SIMILARITY.
FT DISULFID 791 806 BY SIMILARITY.
FT DISULFID 1415 1427 BY SIMILARITY.
FT DISULFID 1417 1434 BY SIMILARITY.
FT DISULFID 1436 1445 BY SIMILARITY.
FT DISULFID 1448 1458 BY SIMILARITY.
FT DISULFID 1461 1469 BY SIMILARITY.
FT DISULFID 1463 1476 BY SIMILARITY.
FT DISULFID 1478 1487 BY SIMILARITY.
FT DISULFID 1490 1503 BY SIMILARITY.
FT DISULFID 1506 1520 BY SIMILARITY.
FT DISULFID 1508 1527 BY SIMILARITY.
FT DISULFID 1529 1538 BY SIMILARITY.
FT DISULFID 1541 1551 BY SIMILARITY.
FT DISULFID 1554 1566 BY SIMILARITY.
FT DISULFID 1556 1573 BY SIMILARITY.
FT DISULFID 1575 1584 BY SIMILARITY.
FT DISULFID 1587 1602 BY SIMILARITY.
FT DISULFID 1830 1839 BY SIMILARITY.
FT DISULFID 1832 1846 BY SIMILARITY.
FT DISULFID 1849 1858 BY SIMILARITY.
FT DISULFID 1861 1877 BY SIMILARITY.
FT DISULFID 1880 1894 BY SIMILARITY.
FT DISULFID 1882 1905 BY SIMILARITY.
FT DISULFID 1907 1916 BY SIMILARITY.
FT DISULFID 1919 1934 BY SIMILARITY.
FT DISULFID 1937 1951 BY SIMILARITY.
FT DISULFID 1939 1958 BY SIMILARITY.
FT DISULFID 1961 1970 BY SIMILARITY.
FT DISULFID 1973 1987 BY SIMILARITY.
FT DISULFID 1990 2000 BY SIMILARITY.
FT DISULFID 1992 2007 BY SIMILARITY.
FT DISULFID 2009 2018 BY SIMILARITY.
FT DISULFID 2021 2034 BY SIMILARITY.
FT DISULFID 2037 2048 BY SIMILARITY.
FT DISULFID 2039 2055 BY SIMILARITY.
FT DISULFID 2057 2066 BY SIMILARITY.
FT DISULFID 2069 2081 BY SIMILARITY.
FT DISULFID 2084 2096 BY SIMILARITY.

FT DISULFID 2086 2103 BY SIMILARITY.
FT DISULFID 2105 2114 BY SIMILARITY.
FT DISULFID 2117 2129 BY SIMILARITY.
FT CARBOHYD 121 121 POTENTIAL.
FT CARBOHYD 140 140 POTENTIAL.
FT CARBOHYD 249 249 POTENTIAL.
FT CARBOHYD 351 351 POTENTIAL.
FT CARBOHYD 477 477 POTENTIAL.
FT CARBOHYD 511 511 POTENTIAL.
FT CARBOHYD 530 530 POTENTIAL.
FT CARBOHYD 634 634 POTENTIAL.
FT CARBOHYD 761 761 POTENTIAL.
FT CARBOHYD 1014 1014 POTENTIAL.
FT CARBOHYD 1341 1341 POTENTIAL.
FT CARBOHYD 1705 1705 POTENTIAL.
FT CARBOHYD 1756 1756 POTENTIAL.
FT CARBOHYD 1868 1868 POTENTIAL.
FT CARBOHYD 1944 1944 POTENTIAL.
FT CARBOHYD 1986 1986 POTENTIAL.
FT CARBOHYD 2002 2002 POTENTIAL.
FT CARBOHYD 2159 2159 POTENTIAL.
FT CARBOHYD 2207 2207 POTENTIAL.
FT CARBOHYD 2231 2231 POTENTIAL.
FT CARBOHYD 2235 2235 POTENTIAL.
FT CARBOHYD 2401 2401 POTENTIAL.
FT CARBOHYD 2421 2421 POTENTIAL.
FT CARBOHYD 2487 2487 POTENTIAL.
FT CARBOHYD 2821 2821 POTENTIAL.
FT CARBOHYD 3087 3087 POTENTIAL.
FT CARBOHYD 3242 3242 POTENTIAL.
FT CARBOHYD 3541 3541 POTENTIAL.
SQ SEQUENCE 3672 AA; 404223 MW; BEC3A008 CRC32;

Query Match 5.6%; Score 97; DB 1; Length 3672;
Best Local Similarity 34.1%; Pred. No. 2.74e-02;
Matches 15; Conservative 8; Mismatches 19; Indels 2; Gaps 2;

Db 552 CQVGFYFGDDC-KFCNCDDPMGTGEGVCDDTTCQCLCKEGFAGDKC 594

QY 184 CDLNYGCGCAKFCRPRDSFGHSTCSET-GEIICLTGWQGDYC 226

RESULT 12
ID DLK_MOUSE STANDARD; PRT; 385 AA.
AC Q09163; Q07645;
DT 01-NOV-1995 (REL. 32, CREATED)
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE DELTA-LIKE PROTEIN PRECURSOR (CONTAINS: FETAL ANTIGEN 1) (FAL) (DLK)
DE (PREADIPOCYTE FACTOR 1) (PREF-1) (ADIPOCYTE DIFFERENTIATION INHIBITOR
DE PROTEIN).
GN DLK OR DLK1 OR PREF1.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SWISS; TISSUE=FIBROBLAST;
RX MEDLINE; 93179372.
RA LABORDA J., SAUSVILLE E.A., HOFFMAN T., NOTARIO V.;
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE; 93272313.
RA SMAS C.M., SUL H.S.;
RL CELL 73:725-734(1993).
RN [3]
RP SEQUENCE OF 1-8 FROM N.A.
RX MEDLINE; 94325292.
RA SMAS C.M., GREEN D., SUL H.S.;
RL BIOCHEMISTRY 33:9257-9265(1994).
CC -1- FUNCTION: MAY HAVE A ROLE IN NEUROENDOCRINE DIFFERENTIATION.
CC INHIBITS ADIPOCYTE DIFFERENTIATION.

CC -1- SUBUNIT: MONOMER.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- PTM: N- AND O-GLYCOSYLATED.
CC -1- SIMILARITY: CONTAINS 6 EGF-LIKE DOMAINS.
DR EMBL: 212171; G50717; -
DR EMBL: U15980; G562108; -
DR EMBL: L12721; G309093; -
DR EMBL: S71340; E120487; -
DR MGD; MGI:94900; DKL1.
DR PROSITE; PS00010; ASX_HYDROXYL; 1.
DR PROSITE; PS00022; EGF_1; 5.
DR PROSITE; PS01186; EGF_2; 6.
KM ANTIGEN: TRANSMEMBRANE. EGF-LIKE DOMAIN. REPEAT: GLYCOPROTEIN; SIGNAL.
FT SIGNAL 1 23
FT CHAIN 24 385
FT PEPTIDE 24 305
FT DOMAIN 24 305
FT TRANSMEM 306 329
FT DOMAIN 330 385
FT DOMAIN 22 55
FT DOMAIN 33 86
FT DOMAIN 88 125
FT DOMAIN 127 168
FT DOMAIN 172 208
FT DOMAIN 210 247
FT DISULFID 26 37
FT DISULFID 30 43
FT DISULFID 45 54
FT DISULFID 57 68
FT DISULFID 63 74
FT DISULFID 76 85
FT DISULFID 92 103
FT DISULFID 97 113
FT DISULFID 115 124
FT DISULFID 131 144
FT DISULFID 138 156
FT DISULFID 158 167
FT DISULFID 176 187
FT DISULFID 181 196
FT DISULFID 198 207
FT DISULFID 214 225
FT DISULFID 219 235
FT DISULFID 237 246
FT CARBOHYD 100 100
FT CARBOHYD 295 295
FT CARBOHYD 342 342
FT CONFLICT 250 250
FT CONFLICT 320 385
SQ SEQUENCE 385 AA; 41320 MW; 8630456D CRC32;
(IN REF. 2).
Query Match 5.5%; Score 95; DB 1; Length 385;
Best Local Similarity 30.5%; Pred. No. 5.81e-02;
Matches 18; Conservative 9; Mismatches 30; Indels 2; Gaps 2;
DB 31 DPQGFCEADVNCRCVCHWEGPLCDK-CVTPAGCV-NGCVCKPMQCKGDKGFCEI 87
QY 170 ESQYTSLEDFRVYCDLNYTSGCAKFCRPRDDSGHSTCSETGITLITLQMGDYCHI 228
RESULT 13
ID YL25 CAEEL STANDARD; PRT; 810 AA.
AC P34418:
DT 01-FEB-1994 (REL. 28, CREATED)
DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)
DT 01-JUN-1994 (REL. 29, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL 89.2 KD PROTEIN F42H10.5 IN CHROMOSOME III.
GN F42H10.5
OS CAENORHABDITIS ELEGANS.
OC EUKARYOTA; METAZOA; ACCELLOMATES; NEMATODA; SECERNENTERA; RHABDITIDA.
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE: 94150718
RA WILSON R., AINSKOUGH R., ANDERSON K., BAYNES C., BERKS M.,
RA BONFIELD J., BURTON J., CONNELL M., COPESEY T., COOULSON A.,
RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FRASER A.,
RA FULTON L., GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M.,
RA JOHNSTON L., JONES M., KERSHAW J., KIRSTEN J., LAISTER N.,
RA LARREILLE P., LIGHTNING J., LLOYD C., MORTIMORE B., O'CALLAGHAN M.,
RA PARSONS J., PERCY C., RIEKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
RA SIMS M., SMALDON N., SMITH A., SMITH M., SONNHAMMER E., STADEN R.,
RA SUSTON J., THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K.,
RA WATERSON R., WATSON A., WEINSTOCK L., WILKINSON-SPROAT J.,
RA WOHLDMAN P.;
RL NATURE 368:32-38(1994).
DR EMBL: L08403; G289681; -
DR WORMPEP; F42H10.5; CE00167.
KM HYPOTHETICAL PROTEIN.
SQ SEQUENCE 810 AA; 89154 MW; 611370C4 CRC32;
Query Match 5.4%; Score 94; DB 1; Length 810;
Best Local Similarity 23.4%; Pred. No. 8.43e-02;
Matches 15; Conservative 16; Mismatches 31; Indels 2; Gaps 2;
DB 676 GGCEALAIYGDLPQSIGNSSESK-ENTVNGYPEISSTSVESMFMTGPNQAPL 734
QY 59 GSCKTRFRCLNHYQATIDTTSQCTGYDVTPIIGENVNLTDAPQFNKGTNPLOPF 118
DB 735 SY-W 737
QY 119 SFSW 122
RESULT 14
ID PAC4 HUMAN STANDARD; PRT; 969 AA.
AC P29122:
DT 01-DEC-1992 (REL. 24, CREATED)
DT 01-DEC-1992 (REL. 24, LAST SEQUENCE UPDATE)
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE SUBTILISIN-LIKE PROTEASE PACE4 PRECURSOR (EC 3.4.21.-).
OS PACE4.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 92075167.
RA KIEFER M.C., TUCKER J.E., JOH R., LANDSBERG K.E., SALTMAN D.,
RA BARR P.J.;
RL DNA CELL BIOL. 10:757-769(1991).
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- TISSUE SPECIFICITY: WIDESPREAD, WITH COMPARATIVELY HIGHER LEVELS
CC IN THE LAYER. PACE4.1 WAS ONLY FOUND IN THE LAYER.
CC -1- ALTERNATIVE PRODUCTS: TWO VARIANTS SEEM TO BE PRODUCED BY
CC ALTERNATIVE SPLICING OF A SINGLE GENE. THEY DIFFER AFTER
CC LVS-471.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8; ALSO KNOWN AS THE
CC SUBTILISIN FAMILY. HIGH SIMILARITY WITH OTHER FURIN-LIKE ENZYMES.
DR EMBL: M80482; G189532; -
DR PIR: A39490; A39490.
DR HSSP; 099405; IMPT.
DR MIM; 167405; -
DR PROSITE; PS00136; SUBTILASE ASP; 1.
DR PROSITE; PS00137; SUBTILASE HTS; 1.
DR PROSITE; PS00138; SUBTILASE SER; 1.
KM HYDROLASE; SERINE PROTEASE; GLYCOPROTEIN; ZIMOGEN; SIGNAL;
KM ALTERNATIVE SPLICING; REPEAT.
FT SIGNAL 1 63
FT PROPEP 64 149
FT CHAIN 150 969
FT DOMAIN 695 969
FT ACT_SITE 205 205
FT ACT_SITE 246 246
FT ACT_SITE 246 246
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).

FT ACT SITE 420 420 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT SITE 553 555 CELL ATTACHMENT SITE (POTENTIAL).
FT DOMAIN 686 941 5 X TANDEM REPEATS, CYS-RICH.
FT REPEAT 686 739 1.
FT REPEAT 740 790 2.
FT REPEAT 791 838 3.
FT REPEAT 839 887 4.
FT REPEAT 888 941 5.
FT CARBOHYD 259 259 POTENTIAL.
FT CARBOHYD 914 914 POTENTIAL.
FT CARBOHYD 932 932 POTENTIAL.
FT VARSPLIC 471 471 K -> KGAAVAFWWTIGWPNV (IN PACE4.1).
FT VARSPLIC 472 969 MISSING (IN PACE4.1).
SQ SEQUENCE 969 AA; 106419 MW; 1FE74A29 CRC32;

Query Match 5.4%; Score 94; DB 1; Length 969;
Best Local Similarity 38.6%; Pred. No. 8.43e-02;
Matches 17; Conservative 6; Mismatches 19; Indels 2; Gaps 2;

DB 733 CPLGYFGDTAARRCR-RCHK-GCETCSSRAATOCILSCRRGFYHH 774
QY 184 CDLNYGSCAKFCRPRDDSGHSTCSTGETGEIICLTGWOGDYCH 227







RESULT 15
ID CANL_YEAST STANDARD; PRT; 590 AA.
AC P04817;
DT 13-AUG-1987 (REL. 05, CREATED)
DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE ARGININE PERMEASE.
GN CANL OR YEL063C.
OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMYCETES.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 88165106.
RA AHMAD M., BUSSEY H.;
RL CURR. GENET. 10:587-592(1986).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE: 86008235.
RA HOFEMAN W.;
RL J. BIOL. CHEM. 260:11831-11837(1985).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-S288C / AB972;
RA DIETRICH F.S., MULLIGAN J.T., HENNESSEY K.M., ALLEN E., ARAUJO R.,
RA AVILES E., BERNO A., BRENNAN T., CARPENTER J., CHEN E., CHERRY J.M.,
RA CHUNG E., DUNCAN M., GUZMAN E., HARTZELL G., HUNICKE-SMITH S.,
RA HYMAN R., KAYSER A., KOMP C., LASHKARI D., LEW H., LIN D.,
RA MOSDALE D., NAKAHARA K., NAMATH A., NORGREN R., OEFNER P., OH C.,
RA PETER F.X., ROBERTS D., SEHL P., SCHRAMM S., SHOGREN T., SMITH V.,
RA TAYLOR P., WEI Y., YELTON M., BOTSTEIN D., DAVIS R.W.;
RL SUBMITTED (DEC-1994) TO EMBL/GENBANK/DDJ DATA BANKS.
CC -!- FUNCTION: HIGH-AFFINITY PERMEASE FOR ARGININE.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -!- SIMILARITY: BELONGS TO A FAMILY THAT GROUPS MANY AMINO ACID
CC PERMEASES.
DR EMBL: X03784; G3442; -.
DR EMBL: M11724; G171155; -.
DR EMBL: U18795; G603255; -.
DR PIR: A23922; QRBYPR.
DR SGD: L0000213; CANL.
DR PROSITE: PS00218; AMINO-ACID TRANSPORT: TRANSMEMBRANE.
KW TRANSPORT; AMINO-ACID TRANSPORT: TRANSMEMBRANE.
FT DOMAIN 1 92 HYDROPHILIC.
FT TRANSMEM 93 110 POTENTIAL.
FT TRANSMEM 230 254 POTENTIAL.
FT TRANSMEM 324 343 POTENTIAL.
FT TRANSMEM 377 396 POTENTIAL.
FT TRANSMEM 420 437 POTENTIAL.
FT TRANSMEM 451 469 POTENTIAL.

FT TRANSMEM 494 518 POTENTIAL.
FT TRANSMEM 525 548 POTENTIAL.
FT DOMAIN 549 590 HYDROPHILIC.
FT CONFLICT 534 534 I -> V (IN REF. 1).
SQ SEQUENCE 590 AA; 65785 MW; 74DCD8FC CRC32;

Query Match 5.3%; Score 92; DB 1; Length 590;
Best Local Similarity 37.0%; Pred. No. 1.76e-01;
Matches 10; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

DB 157 SORFLSPAFAANGANGYMYWFSWATFAL 183
QY 102 AQRFNKGFTNPQPFPSFSGTFSL 128

Search completed: Fri Jun 12 12:05:13 1998
Job time : 20 secs.

FT SIGNAL 1 25 POTENTIAL.
FT CHAIN 26 832 DELTA PROTEIN.
FT DOMAIN 26 595 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 569 617 POTENTIAL.
FT DOMAIN 618 832 INTRACELLULAR (POTENTIAL).
FT DOMAIN 217 566 9 EGF-TYPE REPEATS.
FT REPEAT 217 256 EGF-LIKE 1.
FT REPEAT 257 290 EGF-LIKE 2.
FT REPEAT 291 330 EGF-LIKE 3.
FT REPEAT 331 373 EGF-LIKE 4.
FT REPEAT 374 417 EGF-LIKE 5.
FT REPEAT 418 452 EGF-LIKE 6.
FT REPEAT 453 490 EGF-LIKE 7.
FT REPEAT 491 528 EGF-LIKE 8.
FT REPEAT 529 566 EGF-LIKE 9.
FT CARBOHYD 98 98 POTENTIAL.
FT CARBOHYD 137 137 POTENTIAL.
FT CARBOHYD 167 167 POTENTIAL.
FT CARBOHYD 421 421 POTENTIAL.
FT CARBOHYD 649 649 POTENTIAL.
FT CONFLICT 437 438 GK -> ET (IN REF. 2).
FT CONFLICT 443 443 A -> S (IN REF. 2).
FT CONFLICT 459 459 G -> A (IN REF. 2).
FT CONFLICT 490 490 S -> T (IN REF. 2).
FT CONFLICT 591 591 T -> A (IN REF. 2).
SQ SEQUENCE 832 AA; 88813 MW; CF9ABEC1 CRC32;
Query Match 100.0%; Score 1734; DB 3; Length 832;
Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 230; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 MHWIKLLTAFCFVIVOVHSGSFELRLKYFSDHGRNCRCCSGSDGATGKCLGS 60
Qy 1 MHWIKLLTAFCFVIVOVHSGSFELRLKYFSDHGRNCRCCSGSDGATGKCLGS 60
Db 61 KCTFRVCLKHQATIDTTSQCTGYGVITPILGENSVNLTDQAFONKGFNTNPQPFPSF 120
Qy 61 KCTFRVCLKHQATIDTTSQCTGYGVITPILGENSVNLTDQAFONKGFNTNPQPFPSF 120
Db 121 SWPGTFLIVEAWHDTNNSGNARTNKLIIQRLVQVLEVSSEWTKNSQYTSLEYDF 180
Qy 121 SWPGTFLIVEAWHDTNNSGNARTNKLIIQRLVQVLEVSSEWTKNSQYTSLEYDF 180
Db 181 RVTCDLNYGSCAKFCRPRDDSFHSTCSETGEIICLTGWQGDYCHIPK 230
Qy 181 RVTCDLNYGSCAKFCRPRDDSFHSTCSETGEIICLTGWQGDYCHIPK 230

RESULT 2
ID Q90656 PRELIMINARY; PRT; 728 AA.
AC Q90656;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE TRANSMEMBRANE PROTEIN C-DELTA-1.
OS GALLUS GALLUS (CHICKEN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; AVES; NEOGNATHAE;
OC GALLIFORMES.
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE-SPINAL CORD;
RX MEDLINE; 95319507.
RL NATURE 375:787-790(1995).
DR EMBL; U26590; G882412; -.
DR PROSITE; PS01187; EGF_CA_2.
KW GLYCOPROTEIN; EGF-LIKE DOMAIN.
SQ SEQUENCE 728 AA; 79861 MW; 7439F575 CRC32;
Query Match 44.7%; Score 775; DB 12; Length 728;
Best Local Similarity 49.1%; Pred. No. 1.54e-162;
Matches 104; Conservative 33; Mismatches 74; Indels 1; Gaps 1;

Db 20 QVDGSGVFELKLOEFVNNKGLLSNRNCRGGGPGAGQQQCKTFFRVLCLHQAQSVSP 79
Qy 19 QVHSSGSFELRLKYFSDHGRNCRCCSGSDGATGKCLGSCKTFFRVLCLHQAQSVSP 78
Db 80 EPPCTGSAITPVLGANSFSDGAGADPAFNSPIRFPFGFTWPGTFFSLIIEALHTDSP 139
Qy 79 TSQCTGYGVITPILGENSVNLTDQAFONKGFNTNPQPFPSFWPOTFSLIVEAWH-DTN 137
Db 140 DDLTTPENPERLISRLATQRLAVGEWSQDLHSSGRTDLKYSYRFVCDHYHGGGCVFC 199
Qy 138 NSGNARTNKLIIQRLVQVLEVSSEWTKNSQYTSLEYDFRVTCDLNYGSGCAKFC 197
Db 200 RPRDDFEGHFTGCGEKVCNPGWKGQYCTEP 231
Qy 198 RPRDDFEGHSTCSETGEIICLTGWQGDYCHIP 229

RESULT 3
ID P87357 PRELIMINARY; PRT; 717 AA.
AC P87357;
DT 01-MAY-1997 (TREMBLREL. 03, CREATED)
DT 01-MAY-1997 (TREMBLREL. 03, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE DELTAD TRANSMEMBRANE PROTEIN PRECURSOR.
GN DELTAD.
OS BRACHYDANIO RERIO (ZEBRAFISH) (ZEBRA DANIO).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; PISCES; GNATHOSTOMATA;
OC OSTEICHTHYES; ACTINOPTERYGII; CYPRINIFORMES.
RN [1]
RP SEQUENCE FROM N.A.
RA DORSEIFER P.; TAKKE C.; CAMPOS-ORTEGA J.A.;
RL SUBMITTED (MAR-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; Y11760; E307461; -.
DR PROSITE; PS01187; EGF_CA_2.
KW SIGNAL; TRANSMEMBRANE; GLYCOPROTEIN; EGF-LIKE DOMAIN.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 717 DELTAD TRANSMEMBRANE PROTEIN.
SQ SEQUENCE 717 AA; 79061 MW; 5CC32ECA CRC32;
Query Match 44.3%; Score 769; DB 12; Length 717;
Best Local Similarity 46.0%; Pred. No. 5.29e-161;
Matches 103; Conservative 40; Mismatches 74; Indels 7; Gaps 5;

Db 5 MIAVLIC-VMSIQGFCGVEFELKLOEFVNNKGLLSNRNCRGGGPGAGQQQCKTFFRVLCLHQAQSVSP 79
Qy 7 LTAFAICFTVIVOVHSGSFELRLKYFSDHGRNCRCCSGSDGATGKCLGSCKTFFR 66
Db 61 ICLKHQAQSVSPDCTGYGVITPILGENSVNLTDQAFONKGFNTNPQPFPSFWPOTF 118
Qy 67 VCLKHQAQSVSPDCTGYGVITPILGENSVNLTDQAFONKGFNTNPQPFPSFWPOTF 126
Db 119 SLIIEALHTDSTDDLSTENPDRLISRTTQRHLTVGEWSQDLQVGRTELKYSYRFVCD 178
Qy 127 SLIVEAWH-DTNNSGNARTNKLIIQRLVQVLEVSSEWTKNSQYTSLEYDFRVTCD 185
Db 179 EHYHGGGCVFCRPRDDTFGHFTCGERGEIICNCSHGKQYCTEP 222
Qy 186 LNYGSGCAKFCRPRDDSFHSTCSETGEIICLTGWQGDYCHIP 229

RESULT 4
ID Q00548 PRELIMINARY; PRT; 723 AA.
AC Q00548;
DT 01-JUL-1997 (TREMBLREL. 04, CREATED)
DT 01-JUL-1997 (TREMBLREL. 04, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE DELTA.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
RP SEQUENCE FROM N.A.
RA MANN R.S.; GRAY G.E.; HENRIQUE D.; ISH-HOROWICZ D.,

AC P97607;
DT 01-MAY-1997 (TREMBLREL. 03, CREATED)
DT 01-MAY-1997 (TREMBLREL. 03, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE JAGGED2 (FRAGMENT)
OS RATUUS NORVEGICUS (RAT).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A.
RA SHAWBER C.J., BOUTLER J., LINDSELL C.E., WEINMASTER G.;
RL DEV. BIOL. 180:370-376(1996).
DR EMBL: U70050; G1718248;
DR PROSITE; PS01187; EGF_CA. 7.
KW GLYCOPROTEIN; EGF-LIKE DOMAIN.
FT NON_TER 1
SQ SEQUENCE 1202 AA; 129704 MW; 697F4205 CRC32;

Query Match 28.08; Score 485; DB 10; Length 1202;
Best Local Similarity 40.3%; Pred. No. 2.76e-89;
Matches 79; Conservative 39; Mismatches 64; Indels 14; Gaps 9;

Db 1 GACCDGDRTRAGGCRDCTVYRVCLKEYQAKVTPTGCSYGYGATPVLGNSFVLP 60
QY 43 GRCCSESDEA-TGKC-LGSKTRFRVCLKHQATIDTTSQTYGDTVITPILGENSVNL- 99
Db 61 PAGAAGDRARARSTGGHDPGLVVPFPQFAPRSFTLIVEAW-DWOND-TTPDEELLIE 118
QY 100 -TDA--Q-RFQNK--GFTNP--IQFPFSFPGTSLIVEAWHDTNNSGNARTNKLIIQ 150
Db 119 RVSHAGMINPDRWKSLSFSGVHAHLEQLRVRCDENYSATCNKFCRPRNDFGHYTC 178
QY 151 RLIVQVQVLEVSSEWTKNKSQYTSLEYDFRVTCDLNYGSGCAKFCRPRDDSFHSTCS 210
Db 179 QYGNKACMDGMGKEC 194
QY 211 ETGEIICLTGWQGDYC 226

RESULT 15
ID O35675 PRELIMINARY; PRT; 585 AA.
AC O35675;
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE M-DELTA-LIKE 3 GENE PRECURSOR.
GN M-DELTA-LIKE 3.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-C57B16 X DBA; TISSUE-PRIMITIVE STREAK;
RA DUNWOODIE S.L., HENRIQUE D.M.P., HARRISON S.M., BEDDINGTON R.S.P.;
RL DEVELOPMENT 124:3065-3076(1997).
DR EMBL; Y11895; E322087;
DR MGD; MGI:1096877; MGI:1096877.
DR PROSITE; PS01186; EGF_2; 6.
KW SIGNAL; GLYCOPROTEIN.
FT SIGNAL 1 32 POTENTIAL.
SQ SEQUENCE 585 AA; 61129 MW; 0F5E6187 CRC32;

Query Match 14.4%; Score 249; DB 10; Length 585;
Best Local Similarity 25.3%; Pred. No. 8.14e-33;
Matches 37; Conservative 41; Mismatches 65; Indels 3; Gaps 3;

Db 52 GPCRLFRVCLKPGVSGEATFESICALGALSTSPVYVTEHPGESAAALPLP-DGLVRVPF 110
QY 59 GSKTRFRVCLKHQATIDTTSQTYGDTVITPILGENSVNLDAQRFQNKGTNPQFPF 118
Db 111 RDWPPGTFSLVETREOLGEHAGGPAWNLIA-RVGVRRRLAAGGPWQARDVQRTGTWELH 169
QY 119 SFSWPGTSLIVEAWHDT-NNSGNARTNKLIIQLLVQVQVLEVSSEWTKNKSQYTSLE 177

Db 170 FSYRACEPPAVGAACARLCRSRSAP 195
QY 178 YDFRVTCDLNYYGSGCAKFCRPRDDS 203
Search completed: Fri Jun 12 12:06:25 1998
Job time : 55 secs.

QY	721	 GCCAAGTCTTCGGCGCCGCGACGATTCA	 TTTGACACTCGACTTGC	 TGCAGACGGC	780
Db	781	GAATATATCTGTTGTGACCGGATTA	 CTGTCACATACC	 AAATGCGCCAAA	840
QY	781	GAATATATCTGTTGTGACCGGATTA	 CTGTCACATACC	 AAATGCGCCAAA	840
Db	841	GGCTGTGAACATGGACATTCGACAAACCAATCAATCGTTTGC	 CAACTGCGCAACTGGCTGGAAG	900	
QY	841	GGCTGTGAACATGGACATTCGACAAACCAATCAATCGTTTGC	 CAACTGCGCAACTGGCTGGAAG	900	
Db	901	GGAGCCTTGTGCAACGAGTCGTTCTGGAACCGAACTGCATCCATGGCACCTGCAACAAA	960		
QY	901	GGAGCCTTGTGCAACGAGTCGTTCTGGAACCGAACTGCATCCATGGCACCTGCAACAAA	960		
Db	961	CCCTGGACATTGCATCTGCAACGAGGTTGGGAGGCTTCTACTGCAACACGAGATCTAAC	1020		
QY	961	CCCTGGACATTGCATCTGCAACGAGGTTGGGAGGCTTCTACTGCAACACGAGATCTAAC	1020		
Db	1021	TACTTGCAACCAACACAGACCTGCAAGAATGGCGAACTGCTTCAACACCGCGAGGGA	1080		
QY	1021	TACTTGCAACCAACACAGACCTGCAAGAATGGCGAACTGCTTCAACACCGCGAGGGA	1080		
Db	1081	TTGTACACATGCAAAATGCGCTCCAGGATACAGTGGTGFATGTGGGAAAAATGAGATCTAC	1140		
QY	1081	TTGTACACATGCAAAATGCGCTCCAGGATACAGTGGTGFATGTGGGAAAAATGAGATCTAC	1140		
Db	1141	TCCTGGCATCCGATCTCAATCCCTGCCAGAAATGGTGTACTGTCATCGATGACGCCAC	1200		
QY	1141	TCCTGGCATCCGATCTCAATCCCTGCCAGAAATGGTGTACTGTCATCGATGACGCCAC	1200		
Db	1201	ACAAAACCGGCTACAAGTGTCAATTCGCGCAACGGCTGGAGCGAAAGATGTGCGAGGAG	1260		
QY	1201	ACAAAACCGGCTACAAGTGTCAATTCGCGCAACGGCTGGAGCGAAAGATGTGCGAGGAG	1260		
Db	1261	AAAGTGTCTACGTGTTTCGGACAAAACCTGTCTATCAGGGAATGTCCCGCAAGCTTCGTCTCT	1320		
QY	1261	AAAGTGTCTACGTGTTTCGGACAAAACCTGTCTATCAGGGAATGTCCCGCAAGCTTCGTCTCT	1320		
Db	1321	GGCTTGGGAAGCAAGGTCAGGCTACCACTGCGGAATGTCCATTTGCTACAGCGACCC	1380		
QY	1321	GGCTTGGGAAGCAAGGTCAGGCTACCACTGCGGAATGTCCATTTGCTACAGCGACCC	1380		
Db	1381	AACTGGCATCTCCAGCTGGACAACCTGCAGTCGGAATCCATGATAAOCGGTGAAGTGT	1440		
QY	1381	AACCTGGCATCTCCAGCTGGACAACCTGCAGTCGGAATCCATGATAAOCGGTGAAGTGT	1440		
Db	1441	CAGCCGAGCGGAAGTGTATTGTCGCCAGCGGATTTTGGGAACGAGATGCGAGACCAAC	1500		
QY	1441	CAGCCGAGCGGAAGTGTATTGTCGCCAGCGGATTTTGGGAACGAGATGCGAGACCAAC	1500		
Db	1501	ATTGACGATTGCTTGGCCACCACTGCGAGAACGGACCTGTCATAGATGGTCAAC	1560		
QY	1501	ATTGACGATTGCTTGGCCACCACTGCGAGAACGGACCTGTCATAGATGGTCAAC	1560		
Db	1561	CAATATCGCTGCCAATCGGTTCCCGGTTCCATGCGACCACTGTAGTAGCAAAAGTTGAC	1620		
QY	1561	CAATATCGCTGCCAATCGGTTCCCGGTTCCATGCGACCACTGTAGTAGCAAAAGTTGAC	1620		
Db	1621	TTGTGCCCTCATACAGACCGTGTGCCAATGGAGAACCTGCTTGAATCTCAACACGATTTAC	1680		
QY	1621	TTGTGCCCTCATACAGACCGTGTGCCAATGGAGAACCTGCTTGAATCTCAACACGATTTAC	1680		
Db	1681	CAGTGCACTGTGTCGGGGATTTACTGGCAAGGATTTGCTCTGTGGACATCGATGAGTGC	1740		
QY	1681	CAGTGCACTGTGTCGGGGATTTACTGGCAAGGATTTGCTCTGTGGACATCGATGAGTGC	1740		
Db	1741	AGCAGTGACCCCTGTGCATACCGCGGCACTTGCATGACCCGGTCAATTCGTTTGAATGTC	1800		
QY	1741	AGCAGTGACCCCTGTGCATACCGCGGCACTTGCATGACCCGGTCAATTCGTTTGAATGTC	1800		
Db	1801	GTCGTGCCAATGTTTCAGGGGCAAGCAGTCGATGAGGAGTCTCAGGATTCGGTGACC	1860		

RESULT 3
LOCUS DMDDELTA 2889 bp RNA INV 30-NOV-1994
DEFINITION D.melanogaster DL mRNA for delta protein, involved in neurogenesis.
ACCESSION Y00222
NID 97835
KEYWORDS delta gene; neurogenesis.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Drosophila: Eukaryotes; Metazoa; Arthropoda;
Tracheata: Insecta; Pterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea: Drosophilidae; Drosophila.
1 (bases 6 to 2884)
Kopczynski, C.C., Altom, A.K., Fechtel, K., Kooh, P.J. and
Muskavitch, M.A.
REFERENCE
AUTHORS Delta, a Drosophila neurogenic gene, is transcriptionally complex
and encodes a protein related to blood coagulation factors and
epidermal growth factor of vertebrates
Genes Dev. 2 (12B), 1723-1735 (1988)
JOURNAL 89196890
MEDLINE 2 (bases 1 to 2889)
AUTHORS Muskavitch, M.A.T.
TITLE Direct Submission
JOURNAL Submitted (14-JUN-1991) M.A.T. Muskavitch, Dept. of Biology,
Indiana University, Bloomington, Jordan Hall, A504, Indiana 47405,
USA
COMMENT See X06289 for overlapping sequence.
FEATURES
source
1..2889
/location="Qualifiers"
/organism="Drosophila melanogaster"
/strain="Oregon-R"
/db_xref="taxon:7227"
/haplotype="wildtype"
/dev_stage="4-7 hour embryo"
1..2889
/gene="DL"
/db_xref="FlyBase:FBgn0000463"
/evidence="experimental"
1..2889
/note="Delta"
/gene="DL"
/allele="2"
/db_xref="FlyBase:FBgn0000463"
142..2640
/gene="DL"
/codon_start=1
/product="delta protein"
/db_xref="PID:957774"
/db_xref="FlyBase:FBgn0000463"
/translation="MHVICKILAFICFTIVIVVHSSGSFELRLKPSNDHGRDNEGR
CCSGSDGATGKCLGSCRTFRVLCIKYQATIDTSCCTGADVTTPKIGNSVLLDA
ORONKGFNPIDPFPSFNMFPFLIVEMHDMNSGNRTNLLIOLRLVLOOLV
SSEMKTKSSQSTSLSEYDPRVYCDLTVYSSGCAKPRRPDDSGHSCETGCTIL
TMQGDYCHLPKCAKGCHEHCDKPNOCVQGLKMKALNCEVLEPNCIHGTCKPPT
CICNEWEGGLYCNODLNTYCNHPCKNKGCTFNTGELTYCKAPGSGDCHEVETS
CDADVPCQNGGTCIDEPHTRTKYCKHANGWSMKEEYVLTCSDBPCHQOLRNRY
PTGSKGQGYCCPCPIGSGPNCIDLDONCSNPCIINGSCOPSGKICPAGFSGTC
ETNIDICIGHGCGGTCIDMVNOYRCOCVPGFHTGSCSKVDCLIRPANGCTCN
LNDYDCTCRAFTGKDCSVIDECSSGPGHNGTCNNRNSPEVCYANFRKQDCE
ESTDYTFDAHQGTATTOARADGLTNQVVLINAVFAMPLVAVIAVCVFCMKRRK
RAQEKDAEARKQNEONAVATMHNHNGSVALLASALGCTGNSGLTFDGGNPII
KNTWDSVNNICSAALAAAAAADADECLMGVYASVADNNNANSDFCYAPLQPARS
OKO LNDPPTLHNRGSPAGSSAKGASGGGPGAPAEGRISVLGEGSYCSQWMPSLAAGV
AGACSSOLMAAASAGTGTGAQOORSVYCGTPH"
142..213
/gene="DL"
/db_xref="FlyBase:FBgn0000463"
214..2640
/gene="DL"
/db_xref="FlyBase:FBgn0000463"
/product="delta protein"
820..1839
repeat_region

misc_feature
1999..2016
/note="cysteine-rich, factor IX/EGF-like repeats"
/gene="DL"
/note="Stop-transfer sequence"
/db_xref="FlyBase:FBgn0000463"
BASE COUNT 739 a 764 c 805 g 581 t
ORIGIN
Query Match 99.3%; Score 2871; DB 15; Length 2889;
Best Local Similarity 99.9%; Pred. No. 0.00e+00;
Matches 2889; Conservative 0; Mismatches 0; Indels 3; Gaps 3;
Db 1 GAATCGGAGGAATATTAACAAACATAAACAATTAAGTGTAGTGGCGACA 60
Qy 1 GAATCGGAGGAATATTAACAAACATAAACAATTAAGTGTAGTGGCGACA 60
Db 61 CACACACACACACACACCCGGGATTTATTAACATAAAGCGACTCAATCCAAAAATCA 120
Qy 61 CACACACACACACACACCCGGGATTTATTAACATAAAGCGACTCAATCCAAAAATCA 120
Db 121 GCACAAAAACATCATTAACATGATTTGATTAATTTTATTAACAGACTTCATTGC 180
Qy 121 GCACAAAAACATCATTAACATGATTTGATTAATTTTATTAACAGACTTCATTGC 180
Db 181 TTCACAGCATCGTCAGAGTTCACAGTTCGCGACCTTGATTTGAGTGGCTGAAGTTC 240
Qy 181 TTCACAGCATCGTCAGAGTTCACAGTTCGCGACCTTGATTTGAGTGGCTGAAGTTC 240
Db 241 AGCAACGATCACGGGCGGAGACAGAGGGTCTGCTCAGCGGGGAGTGGAGCGAGCG 300
Qy 241 AGCAACGATCACGGGCGGAGACAGAGGGTCTGCTCAGCGGGGAGTGGAGCGAGCG 300
Db 301 AGGGCAAGTGGCTGGGACGCTGCAAGACGGGTTCCGCTGCTTAAGCACTACAG 360
Qy 301 AGGGCAAGTGGCTGGGACGCTGCAAGACGGGTTCCGCTGCTTAAGCACTACAG 360
Db 361 GCCACCATGACACACACCTCCAGTGCACATACGGGAGCGATACGCCCATCTCGGC 420
Qy 361 GCCACCATGACACACACCTCCAGTGCACATACGGGAGCGATACGCCCATCTCGGC 420
Db 421 GAGAACTGGTCAATCTGACGAGCGCCGAGGCTTCCAGAACAGGGCTTACAGATCC 480
Qy 421 GAGAACTGGTCAATCTGACGAGCGCCGAGGCTTCCAGAACAGGGCTTACAGATCC 480
Db 481 ATCAAGTCCCTTCTCTCATGCGCGGGTACCTTCTGCTGATCGTGAAGCGCTGG 540
Qy 481 ATCAAGTCCCTTCTCTCATGCGCGGGTACCTTCTGCTGATCGTGAAGCGCTGG 540
Db 541 CATGATAGCAACATAGCGGCAATGCGGAATGCGGAACAAGCTCTCATCAGGACTCTG 600
Qy 541 CATGATAGCAACATAGCGGCAATGCGGAATGCGGAACAAGCTCTCATCAGGACTCTG 600
Db 601 GTGACGAGGTACTGAGAGTGTCTCCGCAATGGAAGAGCAAGTGGGAATCGCAGTAC 660
Qy 601 GTGACGAGGTACTGAGAGTGTCTCCGCAATGGAAGAGCAAGTGGGAATCGCAGTAC 660
Db 661 ACCTGCTGAGTACGATTTCCGTCTCACTGCGATCTCAACTACTAGATCCGGCTGT 720
Qy 661 ACCTGCTGAGTACGATTTCCGTCTCACTGCGATCTCAACTACTAGATCCGGCTGT 720
Db 721 GCGAAGTCTGCGGCGCGCGGAGATTTGAGCACTGAGCTTGTGGAGAGCGGGC 780
Qy 721 GCGAAGTCTGCGGCGCGCGGAGATTTGAGCACTGAGCTTGTGGAGAGCGGGC 780
Db 781 GAAATATCTGTTGACCGGATGGAGGGGCTTCTGTCATACCAATATGCGCCAAA 840
Qy 781 GAAATATCTGTTGACCGGATGGAGGGGCTTCTGTCATACCAATATGCGCCAAA 840
Db 841 GCGTGTGAACATGACATTTGGACAAACCAATCAATCGCTTTGGCAACTGGGCTGAG 900
Qy 841 GCGTGTGAACATGACATTTGGACAAACCAATCAATCGCTTTGGCAACTGGGCTGAG 900
Db 901 GGAGCCTTGTCAACGAGTGGCTTGTGGAACCGAATGATCATGACCTGCAACAAA 960

|||||
QY 901 GGAGCCTGTGCAAGAGTGGCTCTGGAACCGAACTGCATCCATGGCACTGCAACAA 960
Db 961 CCCTGGAGTCTGCATCTGCAACAGAGGGTTGGGAGGCTTGTACTGCAACAGAGATCTGAAC 1020
QY 961 CCCTGGAGTCTGCATCTGCAACAGAGGGTTGGGAGGCTTGTACTGCAACAGAGATCTGAAC 1020
Db 1021 TACTGCACCAACACAGACCCCTGCAAGATGGCGGAACCTGTTCAACACCGCGGAGGA 1080
QY 1021 TACTGCACCAACACAGACCCCTGCAAGATGGCGGAACCTGTTCAACACCGCGGAGGA 1080
Db 1081 TTGTACACATGCAAAATGCGCTCCAGGATACAGTGTGATGTTGGGAAATGAGATCTAC 1140
QY 1081 TTGTACACATGCAAAATGCGCTCCAGGATACAGTGTGATGTTGGGAAATGAGATCTAC 1140
Db 1141 TCCTGCGATGCGGATGTCATCCCTGCCAGATGGTGTACCTGATCGATGAGCGGCAC 1200
QY 1141 TCCTGCGATGCGGATGTCATCCCTGCCAGATGGTGTACCTGATCGATGAGCGGCAC 1200
Db 1201 ACAAACCGGCTACAGTGTCTTGGGACAAACCTGTCTATCAGGGAATCTGCGCAACCTTGTCTCT 1260
QY 1201 ACAAACCGGCTACAGTGTCTTGGGACAAACCTGTCTATCAGGGAATCTGCGCAACCTTGTCTCT 1260
Db 1261 AAAGTGCTCACGTGTTGGGACAAACCTGTCTATCAGGGAATCTGCGCAACCTTGTCTCT 1320
QY 1261 AAAGTGCTCACGTGTTGGGACAAACCTGTCTATCAGGGAATCTGCGCAACCTTGTCTCT 1320
Db 1321 GCGTTGGGAAGCAAGGGTCAGGGCTACCACTGCGAATGTCCTATGGCTTACAGCGACCC 1380
QY 1321 GCGTTGGGAAGCAAGGGTCAGGGCTACCACTGCGAATGTCCTATGGCTTACAGCGACCC 1380
Db 1381 AACTGCGATCTCCAGCTGGACAACTGCACTGCGAATGTCCTATGGCTTACAGCGACCC 1440
QY 1381 AACTGCGATCTCCAGCTGGACAACTGCACTGCGAATGTCCTATGGCTTACAGCGACCC 1440
Db 1441 CAGCGAGCGGAAGTGTATTTGCCAGCGGATTTTCCGGGAACGAGATGCGGACCAAC 1500
QY 1441 CAGCGAGCGGAAGTGTATTTGCCAGCGGATTTTCCGGGAACGAGATGCGGACCAAC 1500
Db 1501 ATTGACGATGTCTTGGCCACAGTGTGCGAAGCGGACCTGCAATAGATAGTGTCAAC 1560
QY 1501 ATTGACGATGTCTTGGCCACAGTGTGCGAAGCGGACCTGCAATAGATAGTGTCAAC 1560
Db 1561 CAATATCGTGCATATGCGTTCCTCCGGTTCCATGCGCACCCACTGTAGTAGCAAGTTGAC 1620
QY 1561 CAATATCGTGCATATGCGTTCCTCCGGTTCCATGCGCACCCACTGTAGTAGCAAGTTGAC 1620
Db 1621 TTGTGCTCATCAGACCGGTGTCGAATGGAGGAACCTGCTTGAATCTCAACACGATTAC 1680
QY 1621 TTGTGCTCATCAGACCGGTGTCGAATGGAGGAACCTGCTTGAATCTCAACACGATTAC 1680
Db 1681 CAGTGCACCTGTCTGCGGGATTTACTTGGCAAGGATTTGCTGTGGACATCGATGAGTGC 1740
QY 1681 CAGTGCACCTGTCTGCGGGATTTACTTGGCAAGGATTTGCTGTGGACATCGATGAGTGC 1740
Db 1741 AGCAGTGACCTGTCTAATGCGGCGGCTTGCATGCAACCGGCTCAATTCGTTGCAATGC 1800
QY 1741 AGCAGTGACCTGTCTAATGCGGCGGCTTGCATGCAACCGGCTCAATTCGTTGCAATGC 1800
Db 1801 GTGTGCGCAATGGTTTTCAGGGCAAGCAGTGCATGAGGAGTCCCTACCATTCGTTGACC 1860
QY 1801 GTGTGCGCAATGGTTTTCAGGGCAAGCAGTGCATGAGGAGTCCCTACCATTCGTTGACC 1860
Db 1861 TTCGATGCCCAATATATGAGGCGACCAACAGCGAGAGCGGATGTTGACCAATGCC 1920
QY 1861 TTCGATGCCCAATATATGAGGCGACCAACAGCGAGAGCGGATGTTGACCAATGCC 1920
Db 1921 CAGGTAGTCTAATTCGCTTTTCTCCGTTGCGATGCCCTTTGTTGCGGCTTATTCGCGGG 1980
QY 1921 CAGGTAGTCTAATTCGCTTTTCTCCGTTGCGATGCCCTTTGTTGCGGCTTATTCGCGGG 1980
Db 1981 TSCGTGGTCTTCTGATGAAGCGCAAGCGTAAAGCGTGCAGGAAAGCAGCGGAG 2040
|||||

QY 1981 TSCGTGGTCTTCTGATGAAGCGCAAGCGTAAAGCGTGTCTCAGGAAAGGACGACGCGAG 2040
Db 2041 GCCAGGAAGCAGAAAGCAGAAATCGGTGGCCCAAAATGCATCAAAATGGCGAGTGGGTG 2100
QY 2041 GCCAGGAAGCAGAAAGCAGAAATCGGTGGCCCAAAATGCATCAAAATGGCGAGTGGGTG 2100
Db 2101 GGTGTAGCTTTTGGCTTTCAGCCTCTCTGGCGGCAAAACTGGCGAGCAACAGCGGCTCTCACC 2160
QY 2101 GGTGTAGCTTTTGGCTTTCAGCCTCTCTGGCGGCAAAACTGGCGAGCAACAGCGGCTCTCACC 2160
Db 2161 TTCGATGGCGCAACCGGAATATCATCAAAACACTTGGGACAAGTTCGCTCAACAACATT 2220
QY 2161 TTCGATGGCGCAACCGGAATATCATCAAAACACTTGGGACAAGTTCGCTCAACAACATT 2220
Db 2221 TGTGCTCAGCAGCAGCAGCGGCGGCGGCGGAGCAGCGGCGGAGTGTCTCATGTAC 2280
QY 2221 TGTGCTCAGCAGCAGCAGCGGCGGCGGCGGAGCAGCGGCGGAGTGTCTCATGTAC 2280
Db 2281 GCGGATATGTGCGCTCGGTGGCGGATACAAACATGCAACATGCAACATTTTGTGGCT 2340
QY 2281 GCGGATATGTGCGCTCGGTGGCGGATACAAACATGCAACATTTTGTGGCT 2340
Db 2341 CCGCTACAAAGAGCCAAAGTCGCAAAAGCAACTCAACACCGATCCACGCTCATGACCCG 2400
QY 2341 CCGCTACAAAGAGCCAAAGTCGCAAAAGCAACTCAACACCGATCCACGCTCATGACCCG 2400
Db 2401 GGTTCGCGGAGGAGCTCAGCAGGAGGAGCTTGGCGGAGGACCGGCGGAGCGGAG 2460
QY 2401 GGTTCGCGGAGGAGCTCAGCAGGAGGAGCTTGGCGGAGGACCGGCGGAGCGGAG 2460
Db 2461 GCAAGAGGATCTGTGTTTTAGCGAGGCTTCTTACTGTAGCGAGCTTGGCCCTCGTTG 2520
QY 2461 GCAAGAGGATCTGTGTTTTAGCGAGGCTTCTTACTGTAGCGAGCTTGGCCCTCGTTG 2520
Db 2521 GCGGCGGCGGAGTGGCGGAGGCTTCTTACTGCCAGCTAATGGCTGACGCTTCGCGCAGCG 2580
QY 2521 GCGGCGGCGGAGTGGCGGAGGCTTCTTACTGCCAGCTAATGGCTGACGCTTCGCGCAGCG 2580
Db 2581 GGCA - CGGA - CGGG - ACGGCGCAACAGCAGATCGTGGTCTGGGCACTCCGCAATAG 2637
QY 2581 GGCA - CGGA - CGGG - ACGGCGCAACAGCAGATCGTGGTCTGGGCACTCCGCAATAG 2637
Db 2638 TAACTCCAAATCCGGAAGGCTCTCTGTAATCCGGAAGTCCGAGAAATCCGATGAGGAGCTG 2697
QY 2638 TAACTCCAAATCCGGAAGGCTCTCTGTAATCCGGAAGTCCGAGAAATCCGATGAGGAGCTG 2697
Db 2641 TAACTCCAAATCCGGAAGGCTCTCTGTAATCCGGAAGTCCGAGAAATCCGATGAGGAGCTG 2700
QY 2641 TAACTCCAAATCCGGAAGGCTCTCTGTAATCCGGAAGTCCGAGAAATCCGATGAGGAGCTG 2700
Db 2698 ACAGCACAATACAAAGAAAGACTGGGTTGGGTTCAAAATGTAGAGAGAGCGCAAAAT 2757
QY 2698 ACAGCACAATACAAAGAAAGACTGGGTTGGGTTCAAAATGTAGAGAGAGCGCAAAAT 2757
Db 2701 ACAGCACAATACAAAGAAAGACTGGGTTGGGTTCAAAATGTAGAGAGAGCGCAAAAT 2760
QY 2701 ACAGCACAATACAAAGAAAGACTGGGTTGGGTTCAAAATGTAGAGAGAGCGCAAAAT 2760
Db 2758 GTTGTGTTGATGAAGCAGTTTGTGCTCAGCAAAATGAAAAATCTCTAAGAGCATA 2817
QY 2758 GTTGTGTTGATGAAGCAGTTTGTGCTCAGCAAAATGAAAAATCTCTAAGAGCATA 2817
Db 2818 ACTCGTAACTCCCTAAAAATTTGTATAGTAATAGCAAAAGCTGACCGCGCTTTC 2877
QY 2818 ACTCGTAACTCCCTAAAAATTTGTATAGTAATAGCAAAAGCTGACCGCGCTTTC 2877
Db 2821 ACTCGTAACTCCCTAAAAATTTGTATAGTAATAGCAAAAGCTGACCGCGCTTTC 2880
QY 2821 ACTCGTAACTCCCTAAAAATTTGTATAGTAATAGCAAAAGCTGACCGCGCTTTC 2880
Db 2878 GATCCCGAATTC 2889
QY 2881 GATCCCGAATTC 2892

RESULT 4
LOCUS DMDLG 4724 bp RNA
DEFINITION Drosophila mRNA for Delta (D1) gene.
ACCESSION X06289
NID 97852
KEYWORDS D1 gene; glycoprotein; neurogenesis; transmembrane protein.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryotes; mitochondrial eukaryotes; Metazoa; Arthropoda;
Tracheata; Insecta; Pterygota; Diptera; Brachycera; Muscomorpha;

DU
01 CCCCCGATTAACACAGGGCGCCTCAATCCAAAAAAGGCATTGAACAACAAATAAGAAGTGTTC

126

1156 GTCAATCCCTGCCAGAAATGGTGGTACCTGCATCGATGACCGGCACACAAAAACCGCTAC 1215
Db 1201 AAGTGTCAATTCGCGCAACGGCTGGAGCGGAAAGATGTGGAGAGAAAGTGCTCAGTGT 1260
QY 1216 AAGTGTCAATTCGCGCAACGGCTGGAGCGGAAAGATGTGGAGAGAAAGTGCTCAGTGT 1275
Db 1261 TCGGACAAACCCCTGTCATCAGGGAATCTCCGCAAGCTTGGTCTGCTTGGGAAACGAG 1320
QY 1276 TCGGACAAACCCCTGTCATCAGGGAATCTCCGCAAGCTTGGTCTGCTTGGGAAACGAG 1335
Db 1321 GGTCAAGGGTACCAAGTCCCAATGCTCCATTTGGCTACAGCGGACCAACATTCGATCTCCAG 1380
QY 1336 GGTCAAGGGTACCAAGTCCCAATGCTCCATTTGGCTACAGCGGACCAACATTCGATCTCCAG 1395
Db 1381 GTGGAACACTGCAGTCCGAATCCCTGCATAAAGCGGTGGAGCTGTCCAGCGGAGGAAAG 1440
QY 1396 GTGGAACACTGCAGTCCGAATCCATCATAAAGCGGTGGAGCTGTCCAGCGGAGGAAAG 1455
Db 1441 TGTATTTGCCCCAGCGGATTTTCGGGAAGAGGTGGAGACCAACATTCGATTTGCCCTT 1500
QY 1456 TGTATTTGCCCCAGCGGATTTTCGGGAAGAGGTGGAGACCAACATTCGATTTGCCCTT 1515
Db 1501 GCCCACCAGTGGAGAACGAGGACCTGCATAGATATGTCACCAATATGCTGCCTCA 1560
QY 1516 GCCCACCAGTGGAGAACGAGGACCTGCATAGATATGTCACCAATATGCTGCCTCA 1575
Db 1561 TCGCTTCCCGGTTTCCATGGCACCCACTGTAGTAGCAAAAGTTGCTGCTCATCAGA 1620
QY 1576 TCGCTTCCCGGTTTCCATGGCACCCACTGTAGTAGCAAAAGTTGCTGCTCATCAGA 1635
Db 1621 CCGTGTGCCAATGGAGAACCTGCTTGAATCTCAACAAAGATTACCAAGTGCACCTGCTGT 1680
QY 1636 CCGTGTGCCAATGGAGAACCTGCTTGAATCTCAACAAAGATTACCAAGTGCACCTGCTGT 1695
Db 1681 GCGGGATTTACTGGCAAGATTCCTCGTGGACATCGATGAGTGCAGCAGTGCACCTGT 1740
QY 1696 GCGGGATTTACTGGCAAGATTCCTCGTGGACATCGATGAGTGCAGCAGTGCACCTGT 1755
Db 1741 CATACGCGCGCACTTGCATGAACCCGCTCAATTCGTTGCAATTCGCTGCTGCAATGGT 1800
QY 1756 CATACGCGCGCACTTGCATGAACCCGCTCAATTCGTTGCAATTCGCTGCTGCAATGGT 1815
Db 1801 TTCAGGGGCAAGAGTGCATGAGGAGTCTGATGATTCGCTGATTCGATGCCACCA 1860
QY 1816 TTCAGGGGCAAGAGTGCATGAGGAGTCTGATGATTCGCTGATTCGATGCCACCA 1875
Db 1861 TATGAGCGCACCAACAGCAGAGCCGATGTTAGCCCAATGCCAGTGCCTTAAT 1920
QY 1876 TATGAGCGCACCAACAGCAGAGCCGATGTTAGCCCAATGCCAGTGCCTTAAT 1935
Db 1921 GCTGTTTCTCCGTTCCGATGCTTGGTGGCGGTTATTGCGCGTGCCTGCTTCTGC 1980
QY 1936 GCTGTTTCTCCGTTCCGATGCTTGGTGGCGGTTATTGCGCGTGCCTGCTTCTGC 1995
Db 1981 ATGACGCGCAAGCTGAAGCTGCTCAGGAAAGAGCAACCGGAGGCCAGGAGCAAG 2040
QY 1996 ATGACGCGCAAGCTGAAGCTGCTCAGGAAAGAGCAACCGGAGGCCAGGAGCAAG 2055
Db 2041 GAACAGAAATGCAGTGCACCAATGCATCAATGGCAGTGCCTGCTGCTTGGCT 2100
QY 2056 GAACAGAAATGCAGTGCACCAATGCATCAATGGCAGTGCCTGCTGCTTGGCT 2115
Db 2101 TCAGCCTCTATGCGCGGCAAACTGGCAGCAACAGCGGTCTCACCTTCGATGGCGCAAC 2160
QY 2116 TCAGCCTCTATGCGCGGCAAACTGGCAGCAACAGCGGTCTCACCTTCGATGGCGCAAC 2175
Db 2161 CCGAATATCATCAAAACACCTGGGACAAAGTCCGTCACCAACATTTGTGCTCAGCAGCA 2220
QY 2176 CCGAATATCATCAAAACACCTGGGACAAAGTCCGTCACCAACATTTGTGCTCAGCAGCA 2235
Db 2221 GCAGCGGCGCGCGCAGCAGCGGAGAGTGTCTCATGTACGCGGATATGTGCC 2280
QY 2236 GCAGCGGCGCGCGCAGCAGCGGAGAGTGTCTCATGTACGCGGATATGTGCC 2295

2281 TCGGTGGCGGATTAACAACAATGCCAACTCAGACTTTTGTGCTCGCTACAAAGAGCC 2340
QY 2296 TCGGTGGCGGATTAACAACAATGCCAACTCAGACTTTTGTGCTCGCTACAAAGAGCC 2355
Db 2341 AAGTCCCAAAGCAACTCAACACCGATCCACGCTCATGCTCGGTTCCGCGGAGGC 2400
QY 2356 AAGTCCCAAAGCAACTCAACACCGATCCACGCTCATGCTCGGTTCCGCGGAGGC 2415
Db 2401 ACGTCCGCCAAGGAGCGTCTGGCGGAGGACCGGAGCGCGGAGGCAAGAGATCTCC 2460
QY 2416 AGCTCAGCCAAAGGAGCGTCTGGCGGAGGACCGGAGCGCGGAGGCAAGAGATCTCT 2475
Db 2461 GTTTTGGCGAGGTTCTTCTACTGTAGCAGGTTGGCCCTCTGTTGGCGGCGGAGTG 2520
QY 2476 GTTTTGGCGAGGTTCTTCTACTGTAGCAGGTTGGCCCTCTGTTGGCGGCGGAGTG 2535
Db 2521 GCCGGAGACTGTTTCATCCAA-CTAATGCTGCAGCTTCGGTAGCGGCA-CGSA-CGGG 2577
QY 2536 GCCGGAG-CCGTGTTTCATCCAGCTAATGGCTGCAGCTTCGGCAGCGGCGAGCGGCGG 2594
Db 2578 -ACGGCGCAACAGCAGCGATCCGTGCTGCGCACTCCGCAATTATGTAAGTCTCAAAAT 2636
QY 2595 GACGCGCAACAGCAGCGATCCGTGCTGCGCACTCCGCAAT-ATGTAAGTCTCAAAAT 2653
Db 2637 CCGGAAGGCTCTCGGAATCCGGAATCCGGAATCCGGAATCCGGAATCCGGAATCCGGAAT 2696
QY 2654 CCGGAAGGCTCTCGGTAAATCCGGAATCCGGAATCCGGAATCCGGAATCCGGAATCCGGAAT 2713
Db 2697 AAGAAAGACTGGTGGTTCAAAATGTGAGAGAGCGCCCAAAATGTGTTGTTGAT 2756
QY 2714 AAGAAAGACTGGTGGTTCAAAATGTGAGAGAGCGCC-AAAATGTGTTGTTGAT 2772
Db 2757 TGAAGCAGTTTACTGCTCAGCAAAATGAAAATCTGTAACAGGCAATCTGTAACATC 2816
QY 2773 TGAAGCAGTTTACTGCTCAGCAAAATGAAAATCTGTAACAGGCAATCTGTAACATC 2832
Db 2817 CTTAAAAATTTGTATAGTAATAGCAAGCTGTGACCCAGCGGTTTCGATC 2868
QY 2833 CTTAAAAATTTGTATAGTAATAGCAAGCTGTGACCCAGCGGTTTCGATC 2884

5
RESULT DMDLEGFH 600 bp RNA INV 04-AUG-1995
LOCUS Drosophila melanogaster mRNA fragment of D1 locus with EGF-like
DEFINITION repeats (EGF= epidermal growth factor).
ACCESSION X05140
NID 97851
KEYWORDS D1 gene; EGF-like sequence; epidermal growth factor; neurogenesis;
unidentified reading frame.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryotes; mitochondrial eukaryotes; Metazoa; Arthropoda;
Tracheata; Insecta; Pterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 600)
AUTHORS Knust,E., Dietrich,U., Tepass,U., Bremer,K.A., Weigel,D., Vassini,H.
and Campos-Ortega,J.A.
TITLE EGF homologous sequences encoded in the genome of Drosophila
JOURNAL melanogaster, and their relation to neurogenic genes
MEDLINE EMBO J. 6 (3), 761-766 (1987)
FEATURES 87218537
source 1. .600
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
misc_feature 1. .447
/gene="D1"
/note="4x repeated unit"
/db_xref="FlyBase:FBgn0000463"
gene 1. .600
/note="Delta"
/gene="D1"

/allele=""
/db_xref="FlyBase:FBgn0000463"
<1..>600
/gene="D1"
/note="URP (200 AA) D1 gene"
/codon_start=1
/db_xref="FlyBase:FBgn0000463"
/db_xref="PID:g929563"
/db_xref="FlyBase:FBgn0000463"
/translation="CSPNCPINGSCOPSEFCIPSGFSGTRCTNIDLCIAHCEG
GTCIDMNOYRCOCVPGFHTGCTCTKYLDIRCANGCTCLINNDYOCYCRAGFTG
KDCSVIDDECSGCHNGTGMNKNVSEFECVANGCFRKGOCDESDYSVFDADHGYA
TTGARDGLANNAVILVAFVSMPLVAVIAVIAVAFVFCMKR"

BASE COUNT 133 a 154 c 172 g 141 t
ORIGIN

Query Match 19.5%; Score 564; DB 15; Length 600;
Best Local Similarity 97.0%; Pred. No. 0.00e+00;
Matches 582; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Db 1 TCGAGTCGGAATCCCTGCATTAACGGTGAAGCTGTACAGCCGAGGAACGTGATTTC 60
1405 TCGAGTCGGAATCCCTGCATTAACGGTGAAGCTGTACAGCCGAGGAACGTGATTTC 1464
Db 61 CCCAGCGGATTTGGGGAACGAGTGGGAGAACCAATTGAGATTCGCTTGGCAGCAG 120
1465 CCAGCGGATTTGGGGAACGAGTGGGAGAACCAATTGAGATTCGCTTGGCAGCAG 1524
Db 121 TCGAGAGACGAGGACCTGATGATATGATCAACATATGCTGCAATGCTTCCC 180
1525 TCGAGAGACGAGGACCTGATGATATGATCAACATATGCTGCAATGCTTCCC 1584
Db 181 GCTTTTCATGGCACCACCTAGTAGAGAAAGTTGACTTGTGCTTCATCAGACCGTGTCC 240
1585 GCTTTTCATGGCACCACCTAGTAGAGAAAGTTGACTTGTGCTTCATCAGACCGTGTCC 1644
Db 241 AATGAGAGAACCTCTGATATCTCAACAAGATTAACAGTGCAGTGTGCGGGGATT 300
1645 AATGAGAGAACCTCTGATATCTCAACAAGATTAACAGTGCAGTGTGCGGGGATT 1704
Db 301 ACTGAGAGAACCTCTGATATCTCAACAAGATTAACAGTGCAGTGTGCGGGGATT 360
1705 ACTGAGAGAACCTCTGATATCTCAACAAGATTAACAGTGCAGTGTGCGGGGATT 1764
Db 361 GGCATTGATGAACCGGCTCAATTCGTTCAATGCTGTGTGCAATGTTTCAGGGGC 420
1765 GGCATTGATGAACCGGCTCAATTCGTTCAATGCTGTGTGCAATGTTTCAGGGGC 1824
Db 421 AAGCAGTCGATGAGAGAGTCTTACGATTCGTTGACCTTCGATGCCCAATATGAGCG 480
1825 AAGCAGTCGATGAGAGAGTCTTACGATTCGTTGACCTTCGATGCCCAATATGAGCG 1884
Db 481 ACCACAAAGGAGAGCGGATGTTAGCCCAATGCCAGTAGTCCATATGCTTTC 540
1885 ACCACAAAGGAGAGCGGATGTTAGCCCAATGCCAGTAGTCCATATGCTTTC 1944
Db 541 TCCGTCGATGAGCTTGTGTTGGGCTTATTCGCGGCTGCTGATGATGAGCG 600
1945 TCCGTCGATGAGCTTGTGTTGGGCTTATTCGCGGCTGCTGATGATGAGCG 2004

LOCUS 6 GGU26590 2187 bp mRNA VRT 09-AUG-1995
DEFINITION Gallus gallus C-Delta-1 mRNA, complete cds.
ACCESSION U26590
NID 9882411
KEYWORDS chicken.
SOURCE Gallus gallus
ORGANISM Eukaryote; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Archosauria; Aves; Neognathae; Galliformes;
Phasianidae; Phasianinae; Gallus.

REFERENCE 1 (bases 1 to 2187)
AUTHORS Henrique D., Adam J., Myat A., Chitniss A., Lewis J. and
Ish-Horowitz D.
TITLE Expression of a Delta homologue in prospective neurons in the chick
JOURNAL Nature 375 (6534), 787-790 (1995)
MEDLINE 95319507
REFERENCE 2 (bases 1 to 2187)
AUTHORS Henrique D.
TITLE Direct Submission
JOURNAL Submitted (09-MAY-1995) Domingos Henrique, Dev. Genetics
Laboratory, Dev. Biol. Unit, Imperial Cancer Research Fund, South
Parks Road, Oxford OX1 3PS, UK

FEATURES
source

CDS

1. 2187
/organism="Gallus gallus"
/db_xref="taxon:9031"
/dev_stage="stage 4-6, (1 day old)"
/tissue_type="spinal cord"
/clone_id="E2 chick embryo library from Dr. Tom Jessell,
Columbia University"
1. 2187
/note="Transmembrane protein with 8 EGF-like repeats in
its extracellular region; similar to Drosophila Delta
protein, encoded by Genbank Accession Number Y00222; Mus
musculus Delta-like 1 protein, encoded by Genbank
Accession Number X80903 and Xenopus laevis X-Delta-1 gene,
encoded by Genbank Accession Number I42229"
/codon_start=1
/function="lateral inhibition, presumed ligand for
Notch-like receptors"
/product="C-Delta-1"
/db_xref="PID:g882412"
/translation="MGGRFLTLALSLRCQVDGSGVPEIKQEPVNRKGLLSNR
NCCRGEGGAGQDCKTFKFLVCLKLYSVSEPPCTTGATTPVLAGNSVDP
GAGADPAFNSPIRPFEPFPGTSLILELHNDSPDLTTEPERLSRLATGRHL
AVGEWSDHLSGRTDLKYSRFYCDHYEGCSYCRDRDFRHFPTGEGNEKY
CNPMKGGVCTEPICLPCDDEOHGCDKRGKCRVGNRGYRDCIRYPCGLHGTG
OPMCOEOMEGGLPCNODLNVCTHHKRCRGATCTNNGSGSTCSRGYSSCEI
EINEDANPCKNNGSCDLDENSYSTCTPGRYGKNCLESLATCADGPRFPNGRCDNP
DGGYSCRPPLGYSFNCERKIDYCSSPCANGAACVDLGSYICCCQAGTGRHCDN
VDDASFPVNGGTGCDGVNDYSCFPGYNGKNSFPVSRCHNPCHNGATCHERSN
RYDCEANGYGLNCOFLPEPQGPVVDTEKYTEGONSQFPIAVACGIIIVLM
LLGCAIYVCYRLVOKRHHQPEACRSETEEMNLNANOREKDISIVIGATQIKTN
KKVPHSNDKSNKRYRPREVDVLYHEIKENDSYEHKCEKCEKCTYSEAEKS
AVOLKSSDTSERKRPDYSYTSKDIQSYVVISBEDECIATVEY"

BASE COUNT 504 a 657 c 625 g 401 t
ORIGIN

Query Match 5.8%; Score 169; DB 17; Length 2187;
Best Local Similarity 61.1%; Pred. No. 3.91e-97;
Matches 590; Conservative 0; Mismatches 367; Indels 9; Gaps 7;

Db 58 CAGGTTGACGGCTCGGGGCTGTTGAGCTGAGCTGACGAGATTTGTCAACAAGAGGG 117
196 CAGGTTGACGAGTTCCGGGACGTTTGAGTTGCGCCTGACGATTTACGACGACATCACGG 255
Db 118 CTGCTCAGCAACCGGACACTGTGCGGGGGGGGGCCCGGAGGGCGCGGAGCAGCAG 177
256 CGGAGCAACGAGGGTGGCTGTCAGCGGGGAGTGGAGAGGAGCGAACGGGTAAGTGGCTTC 315
Db 178 TCGGACTGCAAGACCTTCTTCGCGCTGTGCTGGAAGCACTACACAGCCAGCTTCGCC 237
316 GGCACCTGCAAGACCGGCTTTCGCTGCTTAAGCACTACACAGCCAGCCATTCGACAC 375
Db 238 GAGCGGCGCTGACCTAGGAGGAGCGGACATCACCCCGTCTGGGGGCCAACCTTCACG 297
376 ACCTCCAGAGTACCTAGCGGAGGAGCTGATCAGCCCATTCGGGAGAGTGGGCAAT 435
Db 298 GTCCCGGAGCGGGGGGGCGCGCCGCTTCAGCAACCCATCCGCTTCCCTTC 357
436 CTGACCGAGCGCCAGCGCTTCAGCAACAGGGCTTCAGCAATCCATCAGTTCGCCCTTC 495
Db 358 GGCCTTACCTGGCGCGGACCTTTCGCTCATCAGGAGGCTCTGCAACGACGATCCGCC 417

Db	1178	AACCGAGACCTGACACGACACACCACTTAAGCCCTGCAAGAAAGGACACACCTGACAC	1237
0y	1006	AACCGAGACCTGACACGACACCAACACCAACACCTGCAAGAAAGGACACCTGCTTC	1065
Db	1238	AACACGCGGACGAGGAGCTGACACCTGCTTCGCGGCTGGGTACACAGGTG	1289
0y	1066	AACACGCGGAGGAGCTGACACCTGCTTCGCGGCTGGGTACACAGGTG	1117

RESULT 8

LOCUS DRDELTPAD 2342 bp RNA VRT 17-SEP-1997

DEFINITION D.rerio mRNA for Deltad protein.

ACCESSION Y11760

NID 91888391

KEYWORDS deltax gene.

SOURCE zebrafish.

ORGANISM Danio rerio

REFERENCE 1 (bases 1 to 2342)

AUTHORS Dornseifer, P., Takke, C. and Campos-Ortega, J.A.

TITLE Overexpression of a zebrafish homologue of the Drosophila neurogenic gene Delta perturbs differentiation of primary neurons and somite development

JOURNAL Mech. Dev. 63 (2), 159-171 (1997)

MEDLINE 97346722

REFERENCE 2 (bases 1 to 2342)

AUTHORS Campos-Ortega, J.A.

TITLE Direct Submission

JOURNAL Submitted (11-MAR-1997) J. A. Campos-Ortega, Institut fuer Entwicklungsbiologie der, Universitaet zu Koeln, Gyrhofstr. 17, D-50923, Koeln, FRG

FEATURES

source

Location/Qualifiers

1..2342

/organism="Danio rerio"

/db_xref="taxon:7955"

<1..8

9..2162

/gene="deltad"

9..2162

/gene="deltad"

/codon_start=1

/product="deltad transmembrane protein"

/db_xref="pid:e307461"

/db_xref="pid:91888392"

/translation="MGRIMAVLLVCVMISGFCGSEFELKIQEFLNKQVGNANCCGSAAGHGCCECTFEFLIKHQANVSDPCTVGGATPVLAGSNSFQVPSFSDSSSTNPJPAFGFPMGTFSLIHALHTDSDSTETMPDLISRMTOHRLVGEESQLOVGRFELKTSYRFVDEHNYEGGCSYFCRPRDTGFHFGGEGELICNSGKMGQCTEPICLDCEDBGRFDCRPECKCRGVFSKRYCDDCIRYRGCILHGTQOQWQNCQCGWGLFCNQDLNTCTLHKPCCQNGATCTVTGGSTYSCRPFTGDSCEIEVNECSGSRNGSGSCDELENTYSCTPGPGYGRNELSAMTCADGCFNGHGCADNPDEGTYCQCCMGVAFNCEERKIDHCSSNPSCNDAOCLDVLVSYLCCDEGFTGHCEBNIDECAIYPCQNGGTCGCGSLDYCTCTCPGTYGKNKTSYAVKCLHNPCNHCATCHEMNRVYVACIPQYGGNCCFLPBNPOGAAIVGCADKRVSYEEDDGGFPTAVCAGIILVLYLIGGSYFVYIRLKLQORSQOIDSSELEFMANNLNNSRERDLSVSLIGATQVANKIKKDFQSDGNGEFSRSLVDYVNLVHLEKQEDLGRKEDSERSESVARKEPDLSDSEKRNHLKSSDSEKRTESLCKDRKYOSFVLSEKDECIATVEV"

18..41

/gene="deltad"

2163..>2342

BASE COUNT 629 a 590 c 606 g 517 t

ORIGIN

3'UTR

Query Match 4.0%; Score 115; -DB 17; Length 2342; Best Local Similarity 65.18; Pred. No. 1.72e-57; Matches 313; Conservative 0; Mismatches 162; Indels 6; Gaps 4;

Db 512 GCTGAAGTACATATACAGATTGCTTGTGATGATGACATTAACACGCGGAGCGCTGCGGT 571

0y 666 GCTGAAGTACATATATCCGCTGCTACCTGCGATTCACACTACACGATCGCGCTGTGCCAA 725

Db	572	CTCTGCGCCGCGCGGATGATCTTTGGGCACCTTCACTCGGAGCGCGGGAAT	631
Oy	726	GTTCGCGCGCGCGCGAGATTCATTTGGACACTCGACTTGTCTGGAGCGGCGAAT	785
Db	632	TATTCGCAACTCCGATGGAAGGACACTACTGCACAGAACCAATCTCTTCCGGGCTG	691
Oy	766	TATCTGTTTGACCGGATGCGAGCGGATTAATCTGTCACTATCCCAATG-CGCCAAAGC-	843
Db	692	TGATGAGACCAATGGCTTTTGGACAAACCCGGTGATGCAAAATGACAGATGATTTAG	751
Oy	844	TG-TGA---CATGACATTTGGACAAACCCCAATCAATGCGTTTCCCAATGGGCTGGA	899
Db	732	TGGAAGTACTGTGACGACTGCAATCTGCTTACCAGCGCTGTTCATGCGACCTGCCACA	811
Oy	900	GGGAGCCTTGTGCAACGAGTGTGCTTGGAACCGAATCCATCCATGCGACCTTGCAACA	959
Db	812	GCCCTGGCAATGCAACTGCCAAGAGGGTTGGGGAGCTCTCTTCTATACCAAGATCTCA	871
Oy	960	ACCCGAGCTTGCACTCTCAACGAGGGTTGGGAGGCTTGATCTGCAACGAGATCTAA	1019
Db	872	TTACTGCACATCATCAGCCGCTGCGAGATGAGACCACTTGACACACAGCGGAGG	931
Oy	1020	CTACTGCACCAACCAACACCTCTGCAAGATGCGGAACCTTCTTCAACACCGGAGGG	1079
Db	932	AAGTACACACTGCTGATGACAGACTGCGCTTCAACCGGAGACAGCTGTGATGAGTCA	991
Oy	1080	ATTGTACACATGCAATGCAATGCGCTCCAGATACAGTGTGATGATGCGAAATGAGATCTA	1139
Db	992	C 992	
Oy	1140	C 1140	
RESULT	9		
LOCUS	RN078889	2795 bp	mRNA
DEFINITION	Rattus norvegicus Delta1 mRNA, complete cds.		
ACCESSION	U78889		
NID	g1699045		
KEYWORDS			
SOURCE	Norway rat.		
ORGANISM	Rattus norvegicus		
REFERENCE	1 (bases 1 to 2795)		
AUTHORS	Disibio,G., Hebsht,L., Boulter,J. and Weimaster,G.		
TITLE	cDNA sequence of R. norvegicus (rat) Delta1		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 2795)		
AUTHORS	Disibio,G., Hebsht,L., Boulter,J. and Weimaster,G.		
TITLE	Direct Submission		
JOURNAL	Submitted (19-NOV-1996) Biological Chemistry, UCLA School of Medicine, 10833 Leconte Avenue, CHS 33-257, Los Angeles, CA 90095-1737, USA		
FEATURES	Location/Qualifiers		
SOURCE	1. 2795		
	/organism="Rattus norvegicus"		
	/db_xref="taxon:10116"		
	290..2434		
	/note="DSL 11gand"		
	/codon_start=1		
	/product="Delta1"		
	/db_xref="PDB:g1699046"		
	/translation="MGRRSALALAVSALLQVWSSGVEFLKQLEFYNKRLGLENRC		
	CRGSGPACRTFFRYCLKHQASVSPPECTYGSATVAVLGWDSFLPDAGIDPA		
	ESNDIRPPEFTWGTSTLIIHALHDSPPDIATENPERLIISLTGRHLVGEEMSO		
	DHSSGTDLRSTRFYCDEHYIEGGSYCRPRDDAFGHTTCGEREKMDPEMKQO		
	YCDPILPGCDQDHGYCDKPRBECKVWQGRICDCIIRYPCGLBTQCPWQNCQ		
	EGMGKICNDLNTYQCTHHKPCRNKATGNTQOSYITSCPRGYTGAECLEVEBCAS		
	PCRNGSGCTLEDISYSCCTCPGFEYKVCESLAMTACADGPEFNGRCSDNDGCTCHC		
	PAGSGSGEKKIKIDLCSSPSCNGAKVVDIGSYLCRQHGFSGRVCEDNVDDCASPP		
	CANCGTRDSVNDPSCCTCPGYTCRNLSAVYSRCHNAHPCINATGTCRGRYMECQO		

GYGANCQFLLPEPPDLIAAOGGSPWAVCAVGVVLLVLLGCAAVVVRLLKLO
KHOPPPDCGGTETMNNLANCQREKDVSVIIIGATQIKNTNKKADPHGDGADKSSF
KARYPTVDYNLRDLKGDEATVRDAHSKRDTHKQSQSGVGEERSTSLRGEVPPDRKR
PESVYSTKTIQSVILVSAEKDECVIATEV"

BASE COUNT 632 a 788 c 806 g 569 t
ORIGIN

Query Match 3.8%; Score 111; DB 23; Length 2795;
Best Local Similarity 67.08; Pred. No. 1.29e-54;
Matches 278; Conservative 0; Mismatches 131; Indels 6; Gaps 4;

Db 831 ACTACTATGAGAGGCTGCTCGTGTTCGCCACCGCGGATGATGCTCTTGGCCACT 890
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
QY 701 ACTACTACGGATCGGCTGTGCCAAGTTCGCCGCCCGCGGAGATTCATTGGACACT 760
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
Db 891 TCACCTGCGGGGAGAGAGGAGAGATGTCGACCTGCTGCGTGAAGGCGGACTGCA 950
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
QY 761 CGACTTGCTCGGAGACGGCGGAAATATCTGTTGACGGATGCGAGGCGATTACTGTC 820
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
Db 951 CTGACCCCATTTGCTGCCAGGCTGTGATGACCAACATGATATTGACAAACCGGGG 1010
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
QY 821 ACATACCCAAATG-C-GCCAAAGCTG-TGA---ACATGGACATPGCGACAAACCCAAATC 874
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
Db 1011 AATGCAAGTCAGAGTTGGCTGGCAGGCGCTACTGCGATGAATGATCGATACCCAG 1070
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
QY 875 AATGCGTTGCCAAGCTGGCTGGAGGAGGAGCCCTGTGCAAGAGTGGCTTGGACCGA 934
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
Db 1071 GCTGTCTCCATGTTACTGCGCAGAGCCCTGGCAGTGAATGCGCAGAGGCTGGGGG 1130
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
QY 935 ACTGCATCCATGCACTGCAACAAACCTGGACTTGCATCTGCAACGAGGTTGGGAG 994
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
Db 1131 GCCTCTTCTGCAACAGGATCTGAATCTGCACTCACTCACTAAGCCATGCAAGGACGAG 1190
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
QY 995 GCTTGTACTGCAACAGGATCTGAATCTGCACTCACTCACTCACTCACTCACTCACT 1054
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
Db 1191 CCACCTGCACCAACAGCGGCGAGGAGTACACATGCTCTTGGCCGACCGGGTA 1245
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
QY 1055 GAACCTGCTCAACAGCGGAGGAGTGTACACATGCAATGCGTCCAGGATA 1109
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||

RESULT 10 MWDELTA1 2857 bp RNA ROD 29-SEP-1995
LOCUS M.musculus mRNA for Delta-like 1 protein.
DEFINITION X80903
NID 9806569
KEYWORDS delta-like 1 gene.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 2857)
AUTHORS Bettenhausen,B., de Angelis,M.H., Simon,D., Guenet,J.L. and Gossler,A.
TITLE Transient and restricted expression during mouse embryogenesis of Dll1, a murine gene closely related to Drosophila Delta
JOURNAL Development 121 (8), 2407-2418 (1995)
MEDLINE 95401858
REFERENCE 2 (bases 1 to 2857)
AUTHORS Bettenhausen,B.
TITLE Direct Submission
JOURNAL Submitted (11-AUG-1994) B. Bettenhausen, Max-Delbrueck-Labor, Carl-von-Linnee-Weg 10, 50829 Koeln, FRG
REMARK revised by author 11-MAY-95
FEATURES Location/Qualifiers
source 1..2857
/organism="Mus musculus"
/strain="Balb/c x C57BL/6"
/db_xref="taxon:10090"
/dev_stage="d 10.5 p.c."
/clone_lib="lambda Zap II"
/chromosome="17"
/map="8.5 (EUCIB resource)"

gene
CDS

14..2182
/gene="D111"
14..2182
/gene="D111"
/codon_start=1
/product="DELTA-like 1"
/db_xref="PID:g806570"
/translation="MGRSALALAVVSAICVWSSGVFELKLOEFVFNKGLLGNRNC
CRGSSPPCACRTFFRVCLKHQASVSPPEPTCYGSNTVPLGVDSFSLPDGAGIDPA
FSPNIRPFPGTWPFTSLIEALHTDSDPLATENPERLISLTTQRHLTVGEWSQ
DLHSSRTDLRYSYRFVDEHYEGCVFCRDRDAFHFTCGDRGCMCDPWKGO
YCTDPICLPGCDQHGCKRGKRGVQWQYCDICIRYPCGLHGTQCPQWQNCQ
EGWGLFCNODLAVCTHHRPCRNATCTNTGQSYTSCSRPYGTGANCELEVDGCTCHC
PCNGASCTDLEDSCCTCPGPGVGVKVELSAMTCADGPCFNGRSCDNDPDGCTCHC
PLGFSFNCCKMDLGGSPCSNGAKCYDLGNLYLCRQAGFGRCEDNVDCASPP
CANGCTRDSVNDPSCCTPGYTKNCAPVSRCHAPCHNGTAGCHQGRYMWCECAQ
GYGPNCFLLPPFPFPMYVLDLSERHMSQGGFPFPMVAVGAGVILVLLGLCAAV
VYRLKLOKHQPPPCGGTETMNNLANCQREKDVSVIIIGATQIKNTNKKADPHGD
HGA KSSKSVRYPTVDYNLRDLKGDEATVRDTHSKRDTHKQSQSGSAGEBKIAPTLRG
GEIPDRKRPESVYSTSKDTKYQSVYVLSAEKDECVIATEV"
29..61
sig_peptide
674..1564
repeat_region
/gene="D111"
/rpt_family="EGF-like"
/rpt_unit=674..769
/rpt_unit=770..871
/rpt_unit=872..991
/rpt_unit=992..1005
/rpt_unit=1106..1122
/rpt_unit=1123..1336
/rpt_unit=1337..1450
/rpt_unit=1451..1564
BASE COUNT 656 a 751 c 783 g 667 t
ORIGIN

Query Match 3.7%; Score 108; DB 23; Length 2857;
Best Local Similarity 61.6%; Pred. No. 1.81e-52;
Matches 498; Conservative 0; Mismatches 294; Indels 16; Gaps 13;

Db 175 CTGAGGACCTTCTTTGCGGTATGCTCAAGCACTACACAGCCAGCGTGTCCAGGAGCC 234
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
QY 321 CTGCAAGCGGGTTTCGGCTGCTGCTAAAGCACTACACAGCCACCATCGACACACCTC 380
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
Db 235 ACCCTGACCTACGCGAGTGTGTGACGCCAGTGTGGGTGTGAGTCTCTTCAGCTGCC 294
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
QY 381 CCAGTGACCTACGCGGACGTGATCAGCCCATCTCTCGCGAGAACTCGGTCATCTGAC 440
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
Db 295 TGATGGCGCAGGCATC--GACCCGC--CTTCAGCAACCCCATCGATTCCTCCCTCGGCTT 351
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
QY 441 CGAGCCCGAGCGCTTCCAGAAAGGGCTTCACGAATCCCATCGATTCCTCCCTCTCGTT 500
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
Db 352 CACCTGGCAGGTACCTTCTCTGTATCATTTGAAGCCCTCCATACAGACTCTCCCGATGA 411
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
QY 501 CTCATGGCGGGTACCTTCTCGTGTGATCGTGTGAGCCCTGGCATG-ATACGAACAATAGC 559
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
Db 412 CTTGCAACAGAAACCAACAGAAAGTATCATCGCCGCTGACACACAGAGGACCTCAC 471
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
QY 560 GCAATGCGC-GAACAACA-AGCTCCTCATCCAGCGACTCTTGGTGACAGGACTTGGGA 617
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
Db 472 TGTGGGAGAAGTGGTCTCAGGACCT--TCACAGTAGCGCGGCACAGACCTCCGGTACT 530
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
QY 618 GGTGTCTCCGAATGAAGAGACAAAGTTCGGAATCGCAGTA-CACGTGCTGGAGTACG 676
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
Db 531 CTTACCGGTTTGTGTGTGACGAGCACTACTACGGAAGAGTGTGCTGTCTGTCTGCCGAC 590
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
QY 677 ATTTCGGTGTACCTGCGCATCTCACTACTACGATCGGCTGTGCCAAGTCTTCCGCGC 736
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
Db 591 CTCGGGATGAGCCCTTTGGCCACTTACCTCGGGGACAGAGGGGAGAAAGATGTGCGAC 650
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
QY 737 CCCGCGAGGATTCATTTGGACACTCGACTTGTCTCGGAGACGGCGGCAAAATATCTGTTGA 796
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
Db 651 CTGCTGGAAGGCGCAGTACTG-CACGTACCCCAATCTCTGCGCAGGGGTGTGATGACCA 709
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||

97	CCGATGGCAGGGCGGCGATTACTGTCACTATCCAAA-TG-C-GCCAAAGCTG-TGA---A	849
97	CGGATGGCAGGGCGGCGATTACTGTCACTATCCAAA-TG-C-GCCAAAGCTG-TGA---A	849
710	CATGATACACTGTGCAACAACCGAGGAGTGTCAGAGTGCAGAGTTGCTGGCAGGGCCGCTAC	769
850	CATGACACTTTGGCAACAAACCCCAATCAATGCCGTTTCCCAACTGGGCTGGAAAGGAGCTTG	909
770	TGCGATGATGATCCGATACCCAGTTGTTCACATGGACCTGGCAGACCACTCGGCAG	829
910	TGCAACGAGTGGTGTCTGGAAACCGAATGTCATTCATGGCACTGGCAACAAACCTCGACT	969
830	TGTAATCCCAAGGAGGCTGGGGGGGCGCTTTTCTGCAACCAAGACCTGAACTACTGTACT	889
970	TGCATCTGCAACGAGGCTTTGGGAGGCTTGACTGCAACACGAGATCTGAACTACTGACAC	1029
890	CACCATTAAGCCGTGAGAGATGAGACCACTGCAACCAACGCGCCAGGGAGCTACACA	949
1030	AACCAACAACCCCTGCAGAAATGGCGGAACCTGCTTCAACACCGCGAGGATTTGACACA	1089
950	TGTTCTGCGCGACCTGGGATACAGGTG	977
1090	TGCAATGCGCTCGAGATACAGTGGTG	1117

```

BASE COUNT      787 a      634 c      764 g      822 t
ORIGIN
Query Match      3.5%; Score 100; DB 17; Length 3007;
Best Local Similarity 65.8%; Pred. No. 8,92e-47;
Matches 271; Conservative 0; Mismatches 135; Indels 6; Gaps 3;

Db 683 TACTTCGGCGGAGCGCTTGTTCAGATTACTACGCCGCTCTCTGTGACGACAGCGCTGGCCACTAT 742
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 703 TACTACGAGATCCGGGTGTGCCAAAGTTCGCCGCCGCCCGACGATTCATTGTGGACACTCG 762
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 743 ACCCTCGATGGAGATGGGAGCAAGGAATGTTTGGTGGATGGACGAGGAGACTACTGCTCT 802
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 763 ACTTCTCGGAGACGGCGGGAATTAATCTGTGTTTGGACCGGATGGACGGGATTACTGTCTAC 822
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 803 GACCCCATCTGCTCTTTCAGACTGCAGTGAACGACATGTTATTGCGAGTCTCCCGGAGAG 862
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 823 ATACCAAAATGGCGC--CAAG-GCTGTGAC--ATGACATTTGGACAAACCAATCAAA 876
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 863 TGCAGATGTGCTCTGTGGGTGGCAGGGGCTTCTCTGACCGAGTGGCTTCATTATTCAGGA 922
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 877 TGCGTTTGCCAACTGGGTGGAGGGACCTTGTGCCAAGAGTGTGCTTGCAGAACCGAAC 936
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 923 TGCCGTGATGGAGACCTGTTTCACAGCCATGGGACGTGTGTGTAAGAGAGGTTGGGGAGGC 982
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 937 TGCAATCATGGCACCCTGCACAAACCCTTGACATTGCAATCTGCACAGAGGTTGGGAGGC 996
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 983 CTCTTCTGTAACGAGACCTCACTACTGTACCAATCAACAAACCCTGGCTAATGGGCC 1042
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 997 TTGTACTGCAACACGAGATCTGAACTACTGACCAACACAGACACCTGCAAGATGGCGGA 1056
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1043 ACCTGACCAACACCTTGACAGGGCAGCTACACTTGCACCTCGCGCGGGAT 1094
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1057 ACCTGCTTCAACACCGCGGAGGATTTGTACACATGCAAAATGCTCCAGGAT 1108
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 12
LOCUS XELXDEL 2166 bp DNA VRT 17-JUL-1995
DEFINITION Xenopus laevis X-Delta-1 gene, complete cds.
ACCESSION L42229
NID 9807695
KEYWORDS Delta; Notch; X-Delta-1; neurogenic gene.
SOURCE Xenopus laevis DNA.
ORGANISM Xenopus laevis
          Eukaryote; mitochondrial eukaryotes; Metazoa; Chordata;
          Vertebrata; Amphibia; Batrachia; Anura; Mesobatrachia; Pipiloidea;
          Pipidae; Xenopodinae; Xenopus.
          1 (bases 1 to 2166)
          Henrique,D., Adam,J., Myat,A., Chitnis,A., Lewis,J. and
          Ish-Horowicz,D.
          Expression of a Delta homologue in prospective neurons in the chick
          embryo.
          Nature 375 (6534), 787-790 (1995)
          95319507
TITLE 2 (bases 1 to 2166)
AUTHORS Chitnis,A., Henrique,D., Lewis,J., Ish-Horowicz,D. and Kintner,C.
REFERENCE Primary neurogenesis in Xenopus embryos regulated by a homologue of
          the Drosophila neurogenic gene Delta
          Nature 375 (6534), 761-766 (1995)
JOURNAL Nature 375 (6534), 761-766 (1995)
MEDLINE 95319503
REFERENCE 3 (bases 1 to 2166)
AUTHORS Ish-Horowicz,D.
TITLE Direct Submission
JOURNAL Submitted (15-MAY-1995) David Ish-Horowicz, Developmental Biology
          Unit, ICRF, Oxford, UK
FEATURES
          Location/Qualifiers
          1..2166
          /organism="Xenopus laevis"

```

```
/note="(vector lambda gt10)"
/db_xref="taxon:8355"
/clone_lib="Chris Kintner's St17"
1..2166
CDS
    /codon_start=1
    /product="x-Delta-1"
    /db_xref="PID:g807696"
    /translation="MGQRMFLTLVLSAVLCQISCSGLFELRLQEFVNVKGLGNMNC
CRPSLASIQECKTFRRICLKHQSNVSEPPCTYGGAVTPVLGTSFVVPESNA
DPTFSNPRFPFGTPTGLTIIIEAHDSADDLNTENPERLISRLATQRLHTVGEQ
WSQDLHSRDTFLKSYRFVCDYEGGSCDYCRPDADFHFSCGEKGLKCPGW
KGLYCTEPICLPGCDEHGHGDKPGECKRVGWQGRYDECIRYPCGLHGTQQQWQC
NCGEWGLFCNODLNYCTHHKPCENGATCTNGSYTSCRPYGTGSCNCEIEVNEC
DANCKNGSSDLSNYSYSCPPFGYKNCELSANTCAGPCFNGRCADNPQGYI
CFPCVYSGFNEKKIDYCSNPANGARCELDGNSYICQCEGFSGRNCDNDLDCQ
SFPQNGTQCEGIDYSCPTPGYKNCNMPITKCHENPCHNGATCHERNRYVCO
CARGYGNNGOFLPEERPVVDLTERYEGOSGOFPMIACGATVLLMLLGCAY
VVCVRVYQKRHRQPEACRGESKTMNLANQOREKDISVFTGTQIKVNTNKKIDFLS
ESNNEKNGYKPRYPSVDYLVHLEKNEDSPKEERKCAKCSNDSDESDVNSVHSKR
DSSERRRPSAYSSTKDYKQSVVYISDEKDECIATEV"
BASE COUNT      586 a   514 c   541 g   525 t
ORIGIN
    Query Match      3.3%; Score 94; DB 17; Length 2166;
    Best Local Similarity 67.7%; Pred. No. 1,53e-42;
    Matches 180; Conservative 0; Mismatches 86; Indels 0; Gaps 0;
Db 706 CATGATATGTGACAAGCCTGGGGAATGCAATGTAGATGGTGGCAAGCCGCTAC 765
QY 850 CATGGACATTTGGACAACCAATCAATCAATGGTGGCAATGGGCTGGAAGGAGCCTTG 909
Db 766 TGTGATGATGATTCGTTTACCCAGGATCGCTGATGATGATGATGATGATGATGATGAT 825
QY 910 TGAACGAGTGGCTTCTGGAACCGAATGATGATGATGATGATGATGATGATGATGATGAT 969
Db 826 TGCAACTGCAAGAAGCGCTGGGAGGAGCTTTCTGTAACCAAGATCTTAACTACTGACCC 885
QY 970 TGCATCTGCAAGAGGCTGGGAGGAGCTTGTACTGCAACCAAGATCTGAACTACTGACCC 1029
Db 886 CACCAACGACATGCAAAATGAGGACATGCAACCAAGATGCAACCAAGATGCAACCAAGAT 945
QY 1030 AACCAACGACCTTGGCAAGAAGTGGCGAAGCTGTTCAACACCGCGAGGAGTGTACACA 1089
Db 946 TGCTCTGTCGCTGGGTACTGCTGG 971
QY 1090 TGCAAAATGCTCCAGGATACATGG 1115
RESULT 13
LOCUS AF030031 3051 bp mRNA VRT 27-JAN-1998
DEFINITION Danio rerio DeltaA (deltaA) mRNA, complete cds.
ACCESSION AF030031
NID 92809388
KEYWORDS zebrafish.
SOURCE Danio rerio
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Vertebrata; Actinopterygii;
Neopterygii; Teleostei; Euteleostei; Ostariophysi; Cypriniformes;
Cyprinoidae; Cyprinidae; Rasbora; Danio.
REFERENCE 1 (bases 1 to 3051)
AUTHORS Appel,B. and Eisen,J.S.
TITLE Regulation of neuronal specification in the zebrafish spinal cord
by Delta function
JOURNAL Development 125 (3), 371-380 (1998)
MEDLINE 98165392
REFERENCE 2 (bases 1 to 3051)
AUTHORS Appel,B. and Eisen,J.S.
TITLE Direct Submission
JOURNAL Submitted (14-OCT-1997) Institute of Neuroscience, University of
Oregon, 1254 University of Oregon, Eugene, OR 97403, USA
FEATURES
    source
        1..3051
        /organism="Danio rerio"
```

```
/db_xref="taxon:7955"
/chromosome="LG I"
/map="between 21173 and 2450"
1..3051
CDS
    /gene="deltaA"
    /db_xref="PID:g2809389"
    /translation="MNQAKREAFNRDPLTLRSALKLRDFHKLFIYLFIMGRHLL
LILSLILMLCOASSGVFLKLOELFKNGKQVGNKCKGGLTSYQOCCKTFFR
ICLKHVQNASPEPCTYGGTVPVLSNSFOVPTLPDPSFTNIRNFGTWPQTF
SLIICALHADSKEDELTENPERLIISTMTQRLHVGEDWSQDLHVSRYTELYSRFV
DEHYHGGSCSVCRPDADFHFHCGEELICDAGWKQGYCTEPICLPGCDEHGHG
CEKPECKRVGFKGYDECIRYPCGLHGTQQQWQCNGEGLFCNODLNYCTH
HKPCENGATCTNGSYTSCRPFGSGASCEIEVNECTGNPCRGSGCTDMENTYSC
TCPPFGYKNCELSANTCAGPCFNGRCADNPQGYFCQCPYGTGSCNCEIEVNEC
SSPCNGARCDLVNSYLCQCPDGTGMNCDRAGDEC SMYPCQNGTQCEGAGYMTCT
CPPYGTGRNCSFVSRQHNPCNCHATCHERNRYVCAVSGYGRNQCFLLPDRASQ
IASDVPWTAAGSVGLLVLLVAVCAVVVVCVRVQRRRDRREDEVAENETINLIT
NNCHRDKLAVSVGVAVPVKNINKIDFSDHSDLLSTTEKRSYKTRHAPADYNLVHE
VKFEVYKHEVLEHAGKETTMANELSDSCEDIRCSQSDSECTEERKRLKSDASEK
SKYSESRYSESYSKYSKYSKYSKYSKYSKYSKYSKYSKYSKYSKYSKYSKYSKYSKYS
GGKR"
BASE COUNT      798 a   722 c   808 g   723 t
ORIGIN
    Query Match      2.7%; Score 79; DB 17; Length 3051;
    Best Local Similarity 62.1%; Pred. No. 3.92e-32;
    Matches 279; Conservative 0; Mismatches 164; Indels 6; Gaps 4;
Db 916 GCTCAAGTACTCTACCGTTTGTGTCGACGAGCAACTACTATGTTGAAGGATGCTCAGT 975
QY 666 GCTGGAGTACGATTTCCGTGTCACCTCCGATCTCAACTACTAGGATCCGCTGTGCCAA 725
Db 976 GTTCTCCGCGCCAGAGAGCATGCTCGTCACTTCCCTGTCGAGAACGTGGAGAAAT 1035
QY 726 GTTCTCCGCGCCGCGAGAGATTCATTGGACACTCGATTCGTCGAGAGACGGCGAAAT 785
Db 1036 CATCTCGATGCTGGGTGGAAAGGCCAATACTGTCAGAACACCTATCTGCTCGCAGGGT 1095
QY 786 TAICTGTTTACCGGATGCGAGCGGATTAATCTCATATACCCAAATGC--GCCAAG-G 842
Db 1096 CGACGAGGAGCAGGATTTCTGTGAGAAACCTGCGGAGTGCAGTGCAGAGTGGGCTTTAA 1155
QY 843 CTGTGA--A-CATGGACATTGCGACAAACCAATCAATGCGTTTGCCAACTGGCTGGAA 899
Db 1156 AGGCGCTACTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1215
QY 900 GGGAGCTTTGTGCAACGAGTGGCTTCTGGAACCGAACTCATCCATGCGACCTGCAACAA 959
Db 1216 GCCATGCAATGCAACTGCCAGGAAGCTGGGTGGGCTCTCTCTGCAACCAAGATCTGAA 1275
QY 960 ACCTGGACTTGCATCTGCAACGAGGTTGGGAGGCTGTACTGCAACCAAGATCTGAA 1019
Db 1276 CTACTGCACACATCATAGCCCTGTCTGTAATGAGGCCACTTGTAGTAACACGCGCCAGG 1335
QY 1020 CTACTGCACCAACACAGACCCCTGCAAGATGCGGGAACCTGCTTCAACACCGCGGAGG 1079
Db 1336 CAGCTACACCTGTCTCTGTCGCGCCAGAT 1364
QY 1080 ATTGTACATGCAAAATCGCTCCAGGAT 1108
RESULT 14
LOCUS 166494 7218 bp DNA PAT 23-DEC-1997
DEFINITION Sequence 14 from patent US 5670367.
ACCESSION 166494
NID 92724471
KEYWORDS Unknown.
SOURCE
```

```
/organism="Xenopus laevis"  
/db_xref="taxon:8355"  
398. .2326
```

Search completed: Sat Jun 13 08:37:22 1998
Job time : 6237 secs.

Search completed: Sat Jun 13 08:37:22 1998
Job time : 6237 secs.

(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MParch_nun n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Sat Jun 13 09:56:47 1998; MasPar time 377.77 Seconds
973.006 Million cell updates/sec

Tabular output not generated.

Title: >US-08-083-590A-1
Description: (1-2892) from US08083590A.seq
Perfect Score: 2892
N.A. Sequence: 1 GAATTCGGAGCAATTAATCA.....GCCGTTTCGATCCGCAATTC 2892
Comp: CTTAACGCTCTTAATTAATG.....CGCGAAGCTAGGCGCTTAAG

Scoring table: TABLE default
Gap 6

Mmatch STD : Dbase 0; Query 0

Searched: 176886 seqs, 63680241 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: n.geneseq1-2
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part21 22:part22 23:part23
24:part24 25:part25 26:part26 27:part27 28:part28
29:part29 30:part30 31:part31 32:part32 33:part33
34:part34 35:part35 36:part36 37:part37

Statistics: Mean 9.863; Variance 5.746; scale 1.716

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	2888	99.9	2892	5	030997	Delta cDNA D11.	0.00e+00
2	1708	24.5	708	7	043911	Sequence encoding a d	0.00e+00
3	171	5.9	2883	28	F58898	C-Delta-1 gene (alter	1.49e-96
4	147	5.1	2088	28	F58897	C-Delta-1 gene.	1.05e-79
5	131	4.5	2663	34	T70174	Proliferation and dif	1.43e-68
6	108	3.7	2692	28	F58899	M-Delta-1 gene.	9.37e-53
7	97	3.4	1980	28	F59454	H-Delta-1 contig clon	2.71e-45
8	54	1.9	8378	4	Q25811	Drosophila SLIT prote	2.97e-17
9	53	1.8	4208	34	T70175	Proliferation and dif	1.23e-16
10	51	1.8	4483	24	T40091	Human Serrate-2 (H2)	2.09e-15
11	52	1.8	5561	7	Q43910	Sequence encoding a s	5.09e-16
12	53	1.8	6464	24	T40092	Human Serrate-1 (H1)	1.22e-16
13	50	1.7	3579	24	T40092	Chick Serrate cDNA.	8.51e-15
14	45	1.6	204	1	N81164	Base substituted E.co	8.69e-12

Result	ID	Score	Query Match	Length	DB	ID	Description	Pred. No.
15	44	1.5	91	9	051746	Oligonucleotide probe	3.40e-11	
16	44	1.5	91	9	051746	Oligonucleotide probe	3.40e-11	
17	42	1.5	1573	11	068222	Neuroendocrine tumor	5.08e-10	
18	40	1.4	114	12	070466	Generic DNA sequence	7.33e-09	
19	38	1.3	114	12	070469	Generic DNA sequence	1.02e-07	
20	38	1.3	114	12	070468	Generic DNA sequence	1.02e-07	
21	38	1.3	114	12	070465	Generic DNA sequence	1.02e-07	
22	39	1.3	204	1	N81164	Base substituted E.co	2.75e-08	
23	39	1.3	1556	11	068221	Neuroendocrine tumor	2.75e-08	
24	36	1.2	114	12	070467	Generic DNA sequence	1.36e-06	
25	34	1.2	114	12	070470	Generic DNA sequence	1.72e-05	
26	34	1.2	114	12	070472	Generic DNA sequence	1.72e-05	
27	36	1.2	114	12	070469	Generic DNA sequence	1.36e-06	
28	36	1.2	114	12	070468	Generic DNA sequence	1.36e-06	
29	36	1.2	114	12	070467	Generic DNA sequence	1.36e-06	
30	34	1.2	114	12	070466	Generic DNA sequence	1.72e-05	
31	34	1.2	114	12	070465	Generic DNA sequence	1.72e-05	
32	34	1.2	565	6	035072	HCV envelope region n	1.72e-05	
33	33	1.1	39	7	051787	Mixed oligonucleotide	6.02e-05	
34	32	1.1	114	12	070472	Generic DNA sequence	2.08e-04	
35	32	1.1	114	12	070470	Generic DNA sequence	2.08e-04	
36	31	1.1	114	12	070473	Generic DNA sequence	7.06e-04	
37	31	1.1	130	31	T76152	Human vascular cell a	7.06e-04	
38	31	1.1	178	31	T76405	Human endothelin-1 an	2.08e-04	
39	32	1.1	190	31	T76452	Chymase antisense o11	2.08e-04	
40	32	1.1	1089	4	N92576	Sequence of the 1.1kd	2.08e-04	
41	32	1.1	1433	24	Q93243	Insulin like growth f	2.08e-04	
42	32	1.1	1438	1	Q04550	Ecort-Bcort fragment	2.08e-04	
43	32	1.1	4255	23	T37051	Drosophila Fru 2 cDNA	2.08e-04	
44	32	1.1	4835	23	T37050	Drosophila Fru 1 cDNA	2.08e-04	
45	32	1.1	5191	10	Q57710	Neurocan DNA.	2.08e-04	

ALIGNMENTS

RESULT 1
ID 030997 standard; cDNA; 2892 BP.
AC 030997;
DT 01-APR-1993 (first-entry)
DE Delta cDNA D11.
KW Human; Notch; plasmid; cDNA; clone; D11; expression library; PCR;
KW polymerase chain reaction; primer; cloning vector; Delta; Serrate;
KW neurogenic; topolythmic; homotypic; heterotypic; differentiation;
KW quantitation; antibody; ss.
OS Homo sapiens.
FH Key location/Qualifiers
FT cds 142..2643
FT /tag= a
FT /note= "Delta cDNA D11"
→ PD W09219734-A.
PF 12-NOV-1992.
PR 01-MAY-1992: U03651.
PR 03-MAY-1991: US-695189.
PR 14-NOV-1991: US-791923.
PA (INDV) UNIV INDIANA FOUND.
PA (UYA) UNIV YALE.
PI Artavanis-Tsakonas S, Blaumüller CM, Fehon RG, Muskhavitch MAT;
PI Rebay I, Shepard SB;
DR WPI: 92-398861/48.
DR P-PSDB: R28960.
CC Human Notch and Delta DNA and protein sequences - used for study
CC and manipulation of differentiation processes
CC Claim 50; Fig 13; 239PP; English.
CC The sequence given represents the nucleotide sequence of human Delta
CC gene contained in plasmid cDNA clone D11. A human expression library
CC was constructed and screening assays were carried out on to select for
CC the expressed Delta product. Alternatively the sequences could be
CC isolated by amplification using polymerase chain reaction (PCR)
CC primers. The isolated gene may be inserted into a cloning vector and
CC expressed. The Delta gene and also the Notch and Serrate neurogenic
CC genes are designated "topolythmic" genes. The proteins they encode
CC are involved in specific homo- or heterotypic interactions crucial to
CC differentiation. The quantitation of mRNA for human Notch and Delta

CC and adhesive molecules, and study of its expression are possible using
CC the DNA and antibodies raised against the Notch and Delta proteins.
SQ Sequence 2892 BP; 739 A; 763 C; 808 G; 582 T;

Query Match 99.9%; Score 2888; DB 5; Length 2892;
Best Local Similarity 99.9%; Pred. No. 0.00e+00;
Matches 2890; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db	1	gaattcggggaattattcaaacataaaacacaataaacaatttgtagtgcgcaca	60
Qy	1	GAATTCGGAGGAATTATTCAAACATAAACACAATAAACAATTTGAGTAGTTCGCCGACA	60
Db	61	cacacacacacagcccggtgattattacactaaagaagacactcaatcccaaaaatca	120
Qy	61	CACACACACACAGCCCGTGGATTATTACATAAAGCGACACTCAATCCAAAAATCA	120
Db	121	gcaacaaaacatcaataaacatgcatgattgattataatgtttattacagacatttcgc	180
Qy	121	GCACAAAAACATCAATAAACATGCAATGGATTAAATGTTTATTACAGCATTCATTGC	180
Db	181	ttcacagtcatcgtgcaggttcacagttccgagcttttgagttgcctgaagtacttc	240
Qy	181	TTCACAGTCATCGTCAGGTTACAGTTCGGCAGCTTTGAGTTGGCCTGAAAGTACTTC	240
Db	241	agcaacgatacagggcggaacagaggtcgctgctgcaagcgggagtcggagcgagcg	300
Qy	241	AGCAAGGATACGGGCGGACAAACAGAGGTCGCTGTGACGGGGAGTCGGACGGAGCG	300
Db	301	acggggcaagtgcctggcgagtcgaagacgggttcctcgctgctctaaagcactaccag	360
Qy	301	ACGGGCAAGTCCTGGCAGCTGCNAGACGGGTTTCGGCTGTCCTTAAGACTACCAG	360
Db	361	gccacatgcacacacacctccagtcacactacgagggagtgatcaacgccttcctggc	420
Qy	361	GCCACCATGCACACCACTCCAGTCACCTACGGGACGTGATCAACGCCCATCTCGGC	420
Db	421	gagaactcgtcaatctgacagccagccagcttcacagaacagggcttcacgaatccc	480
Qy	421	GAGAACTCGGTCAATCTGACCGAGCGCCAGCGCTTCAGAACAAAGGCTTCACGAATCCC	480
Db	481	atccagttccctctcgttctcatgcccgggtacctctcgtgacgtcgagggccttg	540
Qy	481	ATCCAGTTCCTCTCTCTCTCATGCCGGGTACCTTCTCGTGATCGTAGGCGCTGG	540
Db	541	catgatcacgaacatagcggcaatgcgcgaaccaaagaagctctctatccagcacttg	600
Qy	541	CATGATACGAACAAATAGCGGCAATGCGGAAACCAACAAAGCTCCTCATCCAGCAGCTTG	600
Db	601	gtgcagcaggtactgaggtgtcctccgaatggaagacgaacaaagtgcggaatcgagtac	660
Qy	601	GTGCAGCAGGTACTGGAGGTCTCTCCGAATGGAAGACGAACAAAGTCGGAATCGCAGTAC	660
Db	661	acgtcgtgaggtacgatttcctgacactgcgactcactactacgactcgcggtgt	720
Qy	661	ACGTCGCTGGAGTACGATTCCTGTCACCTGGATCTCAACTACTACGATCCGGCTGT	720
Db	721	gccaaattctgcggcccgagcagatttcattggacactcgcactgctcggagacggcg	780
Qy	721	GCCAAATTCTGCGGCCCGCAGCAGATTCTTTGGACACTCGACTTCTCGGAGACGGCG	780
Db	781	gaaattatctgtttgaccgagtgagcagggcgattactgtcacatacccaaatgcgcaaa	840
Qy	781	GAAATTTATCTGTTTACCGGATGGCAGGCGATTACTGTACATACCAATTCGCCCAAA	840
Db	841	ggctggaacatggacatttcgacaaacgcaatcaatgctttgccaactggcctgaag	900
Qy	841	GGCTGTGAACATGGACATTCGCAAAACCAATCAATGCTTTGCCAACTGGGCTGGAAG	900
Db	901	ggagccttgcgaacgagtgcttctggaacccgaactgcctcatgcacactgcaacaaa	960
Qy	901	GGAGCCTTGTGCAACGAGTCCGTCTGGAAACCGAATGCAATGCATGCCACCTGCAACAAA	960
Db	961	ccctggacttgcatctgcaacagaggtttggggaggtgtgtactgcaacacaggatctgaac	1020

Qy	961	CCCTGACTTGCATCTGCACACGAGGGTTGGGAGGCTTGTACTTGCAACCCAGGATCTGAAC	1020
Db	1021	tactgcaccaaacacagaccctgcaagaaatggcggaacctgtctcaacacccgagggga	1080
Qy	1021	TACTGCACCAACACAGACCCCTGCAGAAATGCGGAACCTGCTTCAACACCGCGAGGGA	1080
Db	1081	ttgtacacatgcataatgcgctccaggatacagtggtgtagattgcgaaaaatgagatctac	1140
Qy	1081	TTGTACACATGCAATGCGCTCCAGGATACAGTGGTGATGATTCGCAAAATGAGATCTAC	1140
Db	1141	tctcgatgcgcatatcaatccctccagaatgggtacctgcactcgatgagcgccac	1200
Qy	1141	TCTCGATGCCATGTCAATCCCTGCCAGAAATGGTGATACCTGCATCGATGAGCGCAC	1200
Db	1201	acaaaacccggtcacagtgatcattgcgcaacggctggagcggaagaatgtgcgaggag	1260
Qy	1201	ACAAAACCCGGCTACAAAGTGTCAATTCGGCCAAACGGCTGGAGCGAAGATGTGCGAGGAG	1260
Db	1261	aaagtgtaacgtgttctggacaaaacctgtcatcaagggaatctgcgcgaacgttcctct	1320
Qy	1261	AAAGTGCTACGTTTCGGACAAACCTGTCTACAGGGAATCTGCCCAACGTTCTGCTCT	1320
Db	1321	ggcttgggaagcaaggtcaggtaccagtcgaaatgtcccatgtgctacagcgaccc	1380
Qy	1321	GGCTTGGGAAGCAAGGTCAGGCTACCACTGCGAATGTCCCATTTGGCTACAGGGACCC	1380
Db	1381	aactgcgactccagctggcaactgcagtcctccaaatccatgcataaaacggtggaactgt	1440
Qy	1381	AACTGCGATCTCCAGCTGGCAACTGCAGTCCGAAATCCATGCATAAACGGTGAAGTGT	1440
Db	1441	cgccgagcggaaggtgtatttgcgaagggatttccgggaacagatcgagacccaac	1500
Qy	1441	CAGCCGAGCGGAAGGTATTTCGCCAGCGGGATTTTCGGGAACGAGATCGAGACCAAC	1500
Db	1501	attgcagattgttggccaccagtcgagaacgagggacacctgcataagatatggtcaac	1560
Qy	1501	ATTGCAGATTGTCTTGGCCACCAGTCGAGACGAGGACCTGCTGTAATCTCAACAACGATTAC	1560
Db	1561	caatatgcgtccaatgcgttcccggttccatggcaccactgtagtagcaaatgtac	1620
Qy	1561	CAATATCGCTGCCAATGCGTTCGCGTTTCCATGGCACCCACCTGTAGTAGCAAGTTGAC	1620
Db	1621	ttgtgcctcatcagaccgctgtgccaatggagggaacctgttgatctcaacaacgattac	1680
Qy	1621	TTGTGCTCATACAGCGCTGCGCAATGAGGAACCTGCTGTAATCTCAACAACGATTAC	1680
Db	1681	cagtgcacctgtcgtcggaatttactggcaagattgctctgtggacatcgatgagtc	1740
Qy	1681	CAGTGCACCTGTCTGCGGATTTACTGGCAAGGATTGCTCTGGGACATCGATGAGTGC	1740
Db	1741	agcagtgcacctgtcataacgcgccgacttcgatgaacccgctcaattcgttcgaatgc	1800
Qy	1741	AGCAGTGCACCTGTCTATAACGGCGGCACCTTGCAATGAACCGGCTCAATTCTGTCGAATGC	1800
Db	1801	gtgtgtcccaatggtttcaggggcaagcagtcgagatgagagtcctcagattcggtaac	1860
Qy	1801	GTGTGTGCCAATGTTTTCAGGGCAAGCAGTCGATGAGGAGTCTTACGATTTCGGTGACC	1860
Db	1861	ttcgatgcccaacaatatgagcgaccacacacgaagcagagccgatggttgaacaaatgcc	1920
Qy	1861	TTCGATGCCCAACCAATATGGAGCGACCAACAGCGAGCGGATGTTTGACCAATGCCC	1920
Db	1921	cagtagtccctaaatgtctgttttccgtgtcgatgctcttggcggttattgcggcg	1980
Qy	1921	CAGTAGTCTTAATGTCTGTTTCTCGTTCGATGCTTGTGGTGGCGGTTATTGCGCGG	1980
Db	1981	tgctggtctctctgcatgaagcgcaagcgttaagcgtctcagggaaaaagcagcgag	2040
Qy	1981	TGCTGCTCTCTCTGCATGAAGCGCAAGCGTAAGCGTGTCTCAGGAAAAAGGACGACGCGAG	2040
Db	2041	gcaaggaagcagaacgaaacgaggttggtggccaacatgcatacaaatggcagtggggtg	2100

QY 2041 GCCAGGAAGCAGAGAACGAGAAATGCGTGGCCACATCATCATCAATGGAGTGGGGTG 2100
DB 2101 ggttagcttggcttcaagctctcttgaggcgcaaaactggcagcaacggcggttcacc 2160
QY 2101 GGTGAGGCTTGGCTTCAGCCCTCTCTGGGGCGGCAAAAGCGAGCAACGCGGTCTCACCC 2160
DB 2161 ttgatggcggcaaccgcaataatcatcaaaaacacctggagcaagtctggttaacaacatt 2220
QY 2161 TTGATGGCGGCAACCCCAATATTCATCAAAAACACCTGGGCAAGTGGTCAACACATTT 2220
DB 2221 tgtgctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2280
QY 2221 TGTGCTCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGTAC 2280
DB 2281 ggcgagatgtgctcgtggtgagcagcagcagcagcagcagcagcagcagcagcagcagc 2340
QY 2281 GCGGATGTGCGCTCGGTGGCGGATTAACAACATGCAACATCAAGCTTTGTGGCT 2340
DB 2341 ccgctacaaagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2400
QY 2341 CCGCTACAAAGAGCCAAATCGCAAAAGCACTCAACACCGATCCACGCTCATGACCCG 2400
DB 2401 ggttcgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2460
QY 2401 GGTTCGCGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2460
DB 2461 ggcagagagatctctgttttagcagcagcagcagcagcagcagcagcagcagcagcagc 2520
QY 2461 GGCAGAGGATCTCGTTTATAGGCGAGGCTTCCTACTGAGCAGCAGTGGCCCTGTTG 2520
DB 2521 ggcgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2580
QY 2521 GCGCGGGGAGGAGTGGCGGAGCCGTTCATCCAGCTAATGGCTGCAGCTTCGGCAGCG 2580
DB 2581 ggcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2640
QY 2581 GGCAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2640
DB 2641 taactccaaaaatccggaagggctctgttaatccggaagaaatccgcaatggagagagctg 2700
QY 2641 TAACTCCAAAAATCCGGAAGGGCTCTGTAAATCCGGGAATCCGCATGAGAGAGCTG 2700
DB 2701 acagacacatacaaaaagaagactgggttggttcaaaatgtgagagagagcagcaaat 2760
QY 2701 ACAGCACATACAAAAGAAAGACTGGGTGGTTCAAAATGTGAGAGAGCGCAAAAT 2760
DB 2761 gttgttgatgagcagctttagctgcaagaaaatgaaaatctgtaacagcagcagcagcagc 2820
QY 2761 GTTGTGTGATGAGAGAGTGTAGTCTCAGGAAAAATGAAAAATCTGTAAACAGGCAAT 2820
DB 2821 actcgttaactccctaaaaaaatctgtatagtaaatagcaagcgtgtgagccagcagcgtt 2880
QY 2821 ACTCGTAAACTCCCTAAAAAATTTGTATAGTAATAGCAAGAGCTGAGCCAGCGTTTC 2880
DB 2881 gatcccgcaatc 2892
QY 2881 GATCCCGCAATTC 2892

RESULT 2
ID Q43911 standard; DNN: 708 BP.

AC Q43911;
DT 30-NOV-1993 (first entry)
DE Sequence encoding a delta protein.
KW Delta; topolythmic protein; family; ss.
OS Drosophila melanogaster.
FH Key Location/Qualifiers
FT cds 1..708
FT /tag- a
PN MO9312141-A.
PD 24-JUN-1993.
PF 11-DEC-1991; U09240.
PR 11-DEC-1991; WO-009240.

PA (UYVA) UNIV YALE.
PI Artavanis-tsakonas S, Fleming RJ;
DR WPI: 93-214095/26.
P-PSDB: R38305.
PT Purified seriate protein, nucleic acid and antibodies - used in
PT the study and manipulation of differentiation and other
PT physiological processes
PS Disclosure: Pages 84-85; 119pp; English.
CC Delta encodes an approx. 100 kd protein (Delta denotes 'DLZM', the
CC protein product of the predominant zygotic and maternal
CC transcripts) that has nine EGF-like repeats within its
CC extracellular domain. Molecular studies have lead to the suggestion
CC that Notch and Delta constitute biochemically interacting elements
CC of a cell communication mechanism involved in early developmental
CC decisions.
SQ Sequence 708 BP; 160 A; 207 C; 192 G; 149 T;
Query Match 24.5%; Score 708; DB 7; Length 708;
Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 708; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 1 atgcattgattaaatgttataaagacatlaattgtctcaagtcacagtcagtcagtcagtcagtc 60
QY 142 ATGCATTGATTAAATGTTTATTAACAGCATTTGCTTCACTCATCGTCGAGTT 201
DB 61 cacagtcggcagcagcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtc 120
QY 202 CACAGTTCGGGAGCTTTAGTTCGCTGAGTACTTACAGCAACATCAGGCGGAGAC 261
DB 121 aacgagagtcgctgctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 180
QY 262 AACGAGGTCGCTGCTGAGCGGGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 321
DB 181 tgcagaagcagtcgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 240
QY 322 TGCAAGCGCGGTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 381
DB 241 cagtcagcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtc 300
QY 382 CAGTGCACCTTACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 441
DB 301 gacgcccagcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtc 360
QY 442 GAGGCCAGCGCTTCAGAAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 501
DB 361 tcatgcccggtaacctctcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtc 420
QY 502 TCATGCGCGGTAACCTTCGCTGATGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 561
DB 421 aatgcggaacacaagaagctctcatcagcagcagtcagtcagtcagtcagtcagtcagtcagtc 480
QY 562 AATGCGCGAACAACAGCTCTCATCCAGCACTCTTGTGTCAGCAGAGGATAGTGGAGGTG 621
DB 481 tctctcgaatggaagcagcagcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtc 540
QY 622 TCTCTCGAATGAGAGAGCAAGTGGATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 681
DB 541 cgtgtaacctgcatcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtc 600
QY 682 CGTGTACCTGCGATCTCACTACTACGATCCGGGTGCGCAAGTTGTGCGGCGCCCGC 741
DB 601 gacgattcatctggaactcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtc 660
QY 742 GACGATTCTTTGAGACTCGACTTCCTGAGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 801
DB 661 tggcagggcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtc 708
QY 802 TGGCAGGCGGATTAAGTCAATACCAATGCGGCAAGGCGTGTGAA 849

RESULT 3
ID T58898 standard; CDNN: 2883 BP.
AC T58898;

28-APR-1997 (first entry)
 C-Delta-1 gene (alternatively spliced variant).
 KW tissue regeneration; Notch; cervix cancer; breast cancer;
 KW lung cancer; colon cancer; melanoma; seminoma;
 KW neurogenesis; therapy; ss.
 OS Gallus sp.
 FH Key Location/Qualifiers
 FT cds 277..2499
 FT /tag= a
 PN WO9701571-A1.
 PD 16-JAN-1997.
 PF 28-JUN-1996; U11178.
 PR 28-JUN-1995; US-000589.
 PA (IMCR) IMPERIAL CANCER RES TECHNOLOGY.
 PI Artavanis-Tsakonas S, Gray GE, Henrique D, Ish-Horowicz D;
 PI Lewis J;
 DR WPI: 97-100159/09.
 DR P-PSDB: W00876.
 DR New vertebrate Delta protein, DNA and antibodies - for treating and
 PT preventing cancer, nervous system disorders and for tissue
 PT regeneration
 PS Disclosure: Fig 1B1-B2; 135pp; English.
 CC The C-delta-1 gene (T58898) codes for the chick homologue (W00876)
 CC of Drosophila Delta, a protein that binds to Notch protein. It was
 CC obtd. by PCR amplification of cDNA from stage 4-6 embryos using
 CC primers (see also T59455-56) based on fly Delta sequences. A
 CC shorter variant (T58897) was also identified. C-Delta-1 expression
 CC appears to be the earliest known marker for prospective neurons.
 CC Mouse (T58899) and human (T58900, T59454) Delta-1 sequences have
 CC also been isolated. Delta-1 genes can be used in the prodn. of
 CC Delta polypeptides and (including antisense sequences) utilised in
 CC the treatment of disorders of cell fate or differentiation, such as
 CC cancer, and nervous system disorders, or to promote tissue
 CC regeneration and repair.
 CC Sequence 2883 BP; 638 A; 858 C; 792 G; 593 T;
 SQ

Query Match 5.9%; Score 171; DB 28; Length 2883;
 Best Local Similarity 61.1%; Pred. No. 1.49e-96;
 Matches 590; Conservative 0; Mismatches 367; Indels 9; Gaps 7;

334 caggttgacggtccgggtgttcgactgaagctgcaggagttgtcaacaagaagggg 393
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 196 CAGGTTACAGTTCGCGACGCTTGAGTTGCGCTGAAGTACTTCACGAACGATCACGGG 255
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 394 ctgctcagcaacgcgaactgtgcggggggggcgcccgagagcgccggcagcagcag 453
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 256 CGGACAAACGAGGGTCCTCTCGAGCGGGGAGTTCGGACGGAGCGGCGCAAGTGCCTG 315
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 454 tgcgactgcaagacctcttcgctgctgctgaagcactaccagggcgctctcccc 513
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 316 GGCAGCTGAAGACGCGGTTTCGGCTCTGCTTAAGCACTACCAGGCGCACCATCGACAC 375
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 514 gagccgcttgcaactcagcagcagcgcacacccccctctcgcgcacactcttcagc 573
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 376 ACCTCCAGTGCACCTACGGGGAGCTGATCACGGCCATTCTCGCGGAGAACTCGGTCAAT 435
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 574 gtcccgacgcgcgggcgccgcagcccgcttcagcaccctcgcgttcccccttc 633
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 436 CTGACGACGCCAGCGCTTCAGAACAAAGGGCTTCAGAACTCCATCCAGTTCCTCCCTTC 495
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 634 ggttcacctggcccgccacctcttcgctcactatcagaggtctgcacaccgactcccc 693
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 496 TCGTTCTATGGCGGGTACTTCTCTGCTGATCTGTCGAGGC-CTGG-CAT-GATACGAAC 552
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 694 gacgacctaccagaaaaaccccgagcgtctatcagcgccttgccaccagagggcac 753
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 553 AATAGGGGCAATGCGGAACCAACAAAGCTCTCTCCAGGCACTCTTGGTTCAGCAGGTA 612
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 754 ctgcggtggcgagagtggtccagagacctgcacagcagcgcccgaccgacctcaag 813
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 613 CTGAGGTGCTCCGGAATGGAGACGAACAAAGTGGGAATCGCAGTACACGTCGCTGGAG 672
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

814 tactctcctcgcttinnngtggatgagcactactacaggggaagctgctgtctcttcgc 873
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 673 TAGGATTTCGGTGTACCTCGGATCTCAACTACTACGGATCCGGCTGTGCCAAGTTCTGC 732
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 874 cggcccggtgacgaccgcttcggtcacttcacctctggtgagagcgtggcgagaaggtctgc 933
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 733 CGGCCCCGCGACGATTTCATTTGGACACTCGACTTGTCTCGGAGACGGGGAAATATCTGT 792
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 934 aaccaggctggaaggccagctactgcaactgagccgattgtgcttgcctgggtgtgaagag 993
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 793 TTGACCGGATGGCAGGGCGATTACTGTACATACCAAAATCGCCAAAGGCTGTGA--A- 849
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 994 cagcagcgtctctgcagaaacctggggaatgcaagtgacagtggtgtgagggggcg 1053
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 850 CATG--GACAT-TGCGACAAACCAATCAATGCGTTTGCCCACTGGGTGGAAGGGAGCC 906
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 1054 tactgtgacgagtgatccgcataccaggtgcctgcacgtactctgcagcagcctac 1113
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 907 TTGTGCAACGAGTGGTCTCGAACCCGAACTGCATCCATGCGACCTGCAACAACCCCTGG 966
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 1114 cagtgcactgccaggaaggtggggcgcccttttctgcaaccagggacctgaactactgc 1173
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 967 ACTTGCATCTCAACGAGGGTGGGGAGGCTTGTACTGCAACAGGATCTGAACCTACTGC 1026
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 1174 actcaccacaagcatgcaagaatgggtccacatgcacacacacacccggtcagggagctac 1233
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 1027 ACCAACACAGACCTCTCAAGATGGCGGAACCTGTTTCAACCGCGGAGGATTTGTAC 1086
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 1234 actgtttcttcgacctggtgcacaggtccagctgcagattgaaatcaacgaatgt 1293
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 1087 ACATGCAATCGCTCCAGGATACAGTGTGATGTTGCGAAATGAGATCTACTCTCTGC 1146
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 1294 gatgcc 1299
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 1147 GATGCC 1152

RESULT 4
 ID T58897 standard; cDNA; 2088 BP.
 AC T58897;
 DT 28-APR-1997 (first entry)
 DE C-Delta-1 gene.
 KW C-Delta-1; cell proliferation; nervous system disorder;
 KW tissue regeneration; Notch; cervix cancer; breast cancer;
 KW lung cancer; colon cancer; melanoma; seminoma;
 KW neurogenesis; therapy; ss.
 OS Gallus sp.
 FH Key Location/Qualifiers
 FT cds 277..2463
 FT /tag= a
 FT /note= "bases 1201-1260 do not encode the
 FT translated sequence for C-Delta-1
 FT polypeptide given in Fig 2, and
 FT bases 1981-2400 are missing from the
 FT sequence"
 PN WO9701571-A1.
 PD 16-JAN-1997.
 PF 28-JUN-1996; U11178.
 PR 28-JUN-1995; US-000589.
 PA (IMCR) IMPERIAL CANCER RES TECHNOLOGY.
 PI Artavanis-Tsakonas S, Gray GE, Henrique D, Ish-Horowicz D;
 PI Lewis J;
 DR WPI: 97-100159/09.
 DR P-PSDB: W11719.
 DR New vertebrate Delta protein, DNA and antibodies - for treating and
 PT preventing cancer, nervous system disorders and for tissue
 PT regeneration
 PS Disclosure: Fig 1A1-1A3; 135pp; English.
 CC The C-delta-1 gene (T58897) codes for the chick homologue (W11719)
 CC of Drosophila Delta, a protein that binds to Notch protein. It was
 CC obtd. by PCR amplification of cDNA from stage 4-6 embryos using

PA (UYVA) UNIV YALE.
 PI Artavanis-Tsakonas S, Gray GE, Henrique D, Ish-Horowicz D;
 PI Lewis J;
 DR WPI: 97-100159/09.
 DR P-PSDB: W11725.
 DR P-PSDB: W11726;
 DR P-PSDB: W11727.
 PT New vertebrate Delta protein, DNA and antibodies - for treating and
 PT preventing cancer, nervous system disorders and for tissue
 PT regeneration
 PS Claim 75: Fig 12B1-6, 135pp; English.
 CC A contig sequence (T59454) of H-Delta-1 gene was obtd. from clone
 CC H0118, which was isolated from foetal brain library using a
 CC partial sequence as probe. Errors in the sequence meant that no
 CC single reading frame gave a correct sequence for the H-Delta-1
 CC protein. Predicted amino acid sequences for all 3 reading frames
 CC (W11725-27) were delt. and sequences (W11728-38) showing homology
 CC to chick and mouse Delta-1 (W11719-20) sequences were identified.
 CC The H-delta-1 gene can be used in the prodn. of H-Delta-1
 CC polypeptides and (including antisense sequences) used to treat
 CC disorders of cell fate or differentiation, such as cancer, and
 CC nervous system disorders, or to promote tissue regeneration and
 CC repair.
 SQ Sequence 1980 BP; 418 A; 591 C; 593 G; 352 T;

Query Match 3.4%; Score 97; DB 28; Length 1980;
 Best Local Similarity 71.9%; Pred. No. 2,71e-45;
 Matches 166; Conservative 0; Mismatches 64; Indels 1; Gaps 1;

Db 422 catgatttctgacaacacccacatgacagtgagtgagccgagc 481
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 850 CATGACATTCGACAAACCAATCAATGCTTGGCAACAGGCTGGAAGGACCTTG 909
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 482 tctgacgagtgatccgctatccagctgtctcattgacagctccagcagcctgagcag 541
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 910 TGCACACGATGCGTTGACACGACATGATGATGACATGACATGACATGACATG 969
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 542 tgcacgtccaggaagantggggggtcttctgcaacagagactgaactactgaca 601
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 970 TGCATCTGCACAGGAGGTGGGAGGCTGTGACACAGGAGCTCACTACTACGAC 1029
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 602 caccataagccctgcaggaatccagcaccctgcacaacacggcgccaggg 652
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 1030 AACCCAGACCCCTGCACAGATGCGGACACTGCTTCAACACCGCGGAGGG 1079
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 8
 ID Q25811 standard; cDNA; 8378 BP.
 AC Q25811:
 DT 05-JAN-1993 (first entry)
 DE Drosophila SLIT protein involved in axon pathway development.
 KW Neurogenesis; BGF-like repeats; epidermal growth factor; TACON;
 KW embryonic CNS; leucine-rich repeat; Flank-LRR-Flank; glial cells;
 KW Notch; axonogenesis; cell-cell interaction; ss.
 OS Drosophila melanogaster.
 FH Key Location/Qualifiers
 FT 5'utr 1..314
 FT cds /*tag= a
 FT 315..4757 /*tag= b
 FT /*product= SLIT_protein
 FT 4755..8378 /*tag= c
 FT 3'utr
 PN W09210518-A.
 PD 25-JUN-1992.
 PF 27-NOV-1991: U09055.
 PR 07-DEC-1990: US-624135.
 PA (UYVA) UNIV YALE.
 PI Artavanis-Tsakonas S, Rothberg JM;
 DR WPI: 92-234590/28.
 DR P-PSDB: R25079.
 PT SLIT protein and sequence elements for treating
 PT neuro-degenerative disease - useful for Alzheimer's disease.

PT nerve damage and Parkinson's disease, for diagnosis of cancer
 PS Claim 2: Page 74-83; 122pp; English.
 CC PCR and standard library screening techniques were used to further
 CC analyse previously isolated slit cDNA clones. A cDNA clone
 CC representing the 5' most 2.4kb of sequence was isolated from a
 CC larval library and PCR was used to isolate a corresp. sequence from
 CC a 4-8 hour embryonic library. Two forms of the SLIT message were
 CC identified differing by 33 nucleotides. Genomic and cDNA sequencing
 CC indicates the transcripts consists of an approximately 314bp 5'
 CC untranslated leader sequence, followed by either a 4407bp or 4440bp
 CC ORF depending on the splice form and a 4kb 3'-UTR.
 SQ Sequence 8378 BP; 2192 A; 2272 C; 2164 G; 1742 T;

Query Match 1.9%; Score 54; DB 4; Length 8378;
 Best Local Similarity 63.9%; Pred. No. 2.97e-17;
 Matches 124; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

Db 3417 tgcgcgaatgagccaagtgcacatgacacattaccacatagctgcatgacgca 3476
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 1525 TGCAGAACGGAGGACCTGCATAGATATGTCACACATATCGCTGCCAATGCGTTCCC 1584
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 3477 ggttcctatgacccaactgcagacatatgacgactgcagaaacacatgtgccag 3536
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 1585 GGTTCATGCGCACCCACTGTAGTACAAAGTTGACTTGCTCATCAGACCGTGTGCC 1644
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 3537 aacgttggaagtggtgtagcagcagcactcaacagcgtgtccagagcactat 3596
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 1645 AATGAGGAACTGCTGTGATCTCAACACGATTACCACTGACCTGTGCGGAGATT 1704
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 3597 acggcgaagtactg 3610
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 1705 ACTGCAAGGATTG 1718
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 9
 ID T70175 standard; cDNA to mRNA; 4208 BP.
 AC T70175:
 DT 11-FEB-1998 (first entry)
 DE Proliferation and differentiation suppression protein encoding cDNA.
 KW Proliferation; differentiation; suppression; human; delta-1;
 KW serrate-1; blood cell; neuron; leukemia; malignant tumour;
 KW immunosuppression; ds.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT CDS 409..4065
 FT /*tag= a
 FT /*product= Differentiation_suppression_protein
 FT 409..501
 FT sig_peptide 410..4062
 FT mat_peptide /*tag= b
 FT /*tag= c
 PN W09719172-A1.
 PD 29-MAY-1997.
 PF 15-NOV-1996: J03356.
 PR 30-NOV-1995: JP-311811.
 PR 17-NOV-1995: JP-299611.
 PA (ASAH) ASAHIT KASEI KOGYO KK.
 PI Itoh A, Sakano S;
 DR WPI: 97-298110/27.
 DR P-PSDB: W18354.
 PT Peptide(s) encoded by human genes delta-1 and serrate-1 - suppress
 PT proliferation and differentiation of undifferentiated human blood
 PT cells
 PS Claim 15: Page 82-91; 114pp; Japanese.
 CC The present sequence encodes a polypeptide which suppresses
 CC proliferation and differentiation of undifferentiated cells such
 CC as neurons and blood cells. The polypeptide may be used for the
 CC prevention and control of disorders involving undifferentiated
 CC cells, such as leukemia and malignant tumours, and improvement of
 CC blood formation, e.g. after immunosuppression.
 SQ Sequence 4208 BP; 1029 A; 1121 C; 1166 G; 892 T;

Query Match 1.8%; Score 53; DB 34; Length 4208;

```
Best Local Similarity 60.1%; Pred. No. 1.23e-16;
Matches 247; Conservative 0; Mismatches 158; Indels 6; Gaps 4;

Db 946 gagtatacagatccggtgacccgctgctgactactactactatgcttggctgcaataagttc 1005
QY 670 GAGTACGATTTCCTCGTCACTACCGGATCTCACTACTACGGATCCGGCTGTGCCAAGTTC 729
Db 1006 tgcgcccccagagatgactcttttgacacactatgctgtgaccagaatggcaacaaaact 1065
QY 730 TGCCGGCCCCGGACGATTTCATTGGACACTCGACTTGCTCGGAGACGGCGGAAATATC 789
Db 1066 tgcattggaaggtggtgagcccccgaatgatacagagctatttgcgcacaagctgcagt 1125
QY 790 TGTTTGACCGGATGGCAGGCGGATTAATCTGTCATATACCACAAATGC-GCCAAAG--GCTGT 846
Db 1126 cctagcagatggtctcttgcgaactcccgagtgactgcaagtgccagtcagcgtggcgaagc 1185
QY 847 G--AA-CATGGACATTGGACAAACCAATCAATGCGTTTGCCAACTGGGCTGGAAGGGA 903
Db 1186 ctgtactgtataagtgatccacacacccgggagtgctccagggcactgtaatgagccc 1245
QY 904 GCGTTGTGCAAGAGTGCCTTCTGGAAACCGAAGTGCATCCATGGCACCTGCAACAAACCC 963
Db 1246 tggcagtgctctgtgagaccactggggcgccagctctgtgacaaagatctcaattac 1305
QY 964 TGGACTTGCATCTGCAACAGAGGTTGGGAGGCTTGTACTGCAACGAGGATCTGAACTAC 1023
Db 1306 tgtgggaactcaagcgtgtctcaacgcggaactgttagcaacacagc 1356
QY 1024 TGCACCAACACAGACCCCTGCAAGATGCGGGAACCTGCTTCAACACCGGC 1074

RESULT 10
ID T40091 standard; cDNA; 4483 BP.
AC T40091;
DE Human Serrate-2 (HJ2) cDNA (containing an internal deletion).
KW Serrate-2; human jagged-2; HJ2; Notch; cell differentiation;
KW cell fate; central nervous system; cancer; tissue repair; therapy;
KW diagnosis; ds.
OS Homo sapiens.
FH Key Location/Qualifiers
FT cds 332..4105
FT misc_difference 240..241
FT /tag= a
FT /note= "there is a deletion of approx. 120
nucleotides in this region of the sequence,
probably resulting from a cloning artifact
in construction of the library, encoding a
portion of the signal peptide and beginning
of the DSL domain"
FT W09627610-A1.
PD 12-SEP-1996.
PN 07-MAR-1996; U03172.
PR 07-MAR-1995; US-400159.
PA (IMCR ) IMPERIAL CANCER RES TECHNOLOGY.
PA (UYA ) UNIV YALE.
PI Artavanis-Tsakonas S, Gray GE, Henrique DMP, Ish-Horowicz D;
PI Lewis JH, Mann RS, Myat AM;
DR WPI: 96-425379/42.
DR P-PSDB; W05834.

PT Vertebrate Serrate protein and related DNA - used to treat or
prevent malignancies characterised by increased Notch activity.
PS Disclosure: Page 98-103; 161pp; English.
CC 2 cDNA clones (T40090 and T40091) respectively code for human
Serrate-1 (W05833) and human Serrate-2 (W05833), ligands for Notch
that are believed to play a major role in determining cell fates in
the central nervous system. The clones were isolated from a foetal
brain cDNA library using probes obtd. by PCR amplification of human
placental cDNA with primers based on conserved sequences of
Drosophila Serrate and Delta. The cDNA can be incorporated into
CC vectors and utilised in the treatment and diagnosis of disorders of
```

```
CC cell fate and differentiation; antisense sequences can be used to
CC treat malignancies of cell types that express Serrate or Notch.
SQ Sequence 4483 BP; 780 A; 1389 C; 1472 G; 842 T;

Query Match 1.8%; Score 51; DB 24; Length 4483;
Best Local Similarity 64.1%; Pred. No. 2.09e-15;
Matches 116; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

Db 620 ggcctggaagggaggttctgcatgagtggtccctaccctggtgctgcatggcagt 679
QY 892 GGCCTGGAAGGGAGCCTTGTGCAACGAGTGCCTTCTGGAACCGAAGTGCATCCATGGCACC 951
Db 680 tggctgagccctggcagtgcaactgtgagacaaactggggcgccctgctctgtgacaaa 739
QY 952 TGCACAAACCCCTGGACTTGCATCTGCACGAGGGTGGGGAGGCTTGTACTGCACACCAG 1011
Db 740 gacctgaactactgtggcagccaccacccctgcacaaaggaggcagctgcatcaacgcc 799
QY 1012 GATCTGAAGTACTGTCACCAACACACAGACCTGCAAGAATGGCGGAACCTGCTTCAACACC 1071
Db 800 g 800
QY 1072 G 1072

RESULT 11
ID Q43910 standard; DNA; 5561 BP.
AC Q43910;
DE 30-NOV-1993 (first entry)
DE Sequence encoding a serrate protein.
KW Serrate; topothymic protein; family; ss.
OS Drosophila melanogaster.
FH Key Location/Qualifiers
FT cds 442..4653
FT /tag= a
FT W09312141-A.
PD 24-JUN-1993.
PN 11-DEC-1991; U09240.
PR 11-DEC-1991; WO-U09240.
PA (UYA ) UNIV YALE.
PI Artavanis-Tsakonas S, Fleming RJ;
DR WPI: 93-214095/26.
DR P-PSDB; R38304.
PT Purified serrate protein, nucleic acid and antibodies - used in
the study and manipulation of differentiation and other
physiological processes
PS Claim 39; Pages 74-80; 119pp; English.
CC Two Drosophila genomic phage libraries were screened and recombinant
clones were isolated. The cDNAs in lambda gt10 were isolated from an
early pupal library. The C1 cDNA was isolated from an early pupal
CC library. Subsequently the C3 cDNA was isolated using the 5' 700 bp
terminal fragment of the C1 cDNA as probe. The complete 5561bp
CC sequence of DNA of the Drosophila Serrate protein was derived from
CC C1 and C3 cDNAs (Q43910). The deduced protein product appears to be
a transmembrane protein. AAs 51-80 represent the likely signal
peptide; aAs 542-564 represent potential membrane associated region;
CC aAs 1221-1245 represent the putative transmembrane domain.
SQ Sequence 5561 BP; 1398 A; 1447 C; 1504 G; 1212 T;

Query Match 1.8%; Score 52; DB 7; Length 5561;
Best Local Similarity 66.7%; Pred. No. 5.09e-16;
Matches 134; Conservative 0; Mismatches 64; Indels 3; Gaps 3;

Db 1340 aatgcgaatgcagaccccgctgctgctcattgtgcaacgagtgctggtctatcccg 1399
QY 875 AATGCGTTTCCCAACTGGGCTGCAAGGAGCCTTGTGCAACGAGTGGTGTCTGGAACCGA 934
Db 1400 gctgcgaagcagtggttctcccaacgagcagcgctggaatgcgtgctgcaacacactgg 1459
QY 935 ACTGCTCATTCATGCGACCTGCAACAA-A-C-CTTGGACTTGTGATCTGTCACAGGGTGGG 991
Db 1460 gtggcattgtgcgataagatttaattctgcggcaccatgaacctgaacgacgacg 1519
```


CC TOPOLOGY: unknown
CC MOLECULE TYPE: CDNA
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 142..2640
SQ Sequence 2892 bp; 739 A; 764 C; 808 G; 581 T; 0 other;

Query Match 100.0%; Score 2892; DB 1; Length 2892;
Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 2892; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 GAATTCGGAGGAATTATTCAAAACATAAACAACAATAAACAATTTGAGTAGTTGCCGCACA 60
Qy 1 GAATTCGGAGGAATTATTCAAAACATAAACAACAATAAACAATTTGAGTAGTTGCCGCACA 60

Db 61 CACACACACACAGCCCGTGGATTATTACACTAAAAGCGACACTCAATCCAAAATAACA 120
Qy 61 CACACACACACAGCCCGTGGATTATTACACTAAAAGCGACACTCAATCCAAAATAACA 120

Db 121 GCACAAAAACATCAATAAATGATGATTTAAATGTTTATTAAACAGCAATTCATTTC 180
Qy 121 GCACAAAAACATCAATAAATGATGATTTAAATGTTTATTAAACAGCAATTCATTTC 180

Db 181 TTCACAGTCAATCGTCAGGTTTCAAGTTTCGGCAGCTTTGAGTTGCGCTGAAGTACTTC 240
Qy 181 TTCACAGTCAATCGTCAGGTTTCAAGTTTCGGCAGCTTTGAGTTGCGCTGAAGTACTTC 240

Db 241 AGCAACGATCACGGGCGGGAACAGAGGTCGCTGCAGCGGGGAGTCGGACGGAGCG 300
Qy 241 AGCAACGATCACGGGCGGGAACAGAGGTCGCTGCAGCGGGGAGTCGGACGGAGCG 300

Db 301 ACGGGCAAGTGCCTGGGCAGCTGCAAGACGGGGTTTCGGCTCTGCCTTAAAGCACTACAC 360
Qy 301 ACGGGCAAGTGCCTGGGCAGCTGCAAGACGGGGTTTCGGCTCTGCCTTAAAGCACTACAC 360

Db 361 GCCACCATGCACACACCTCCAGTGCACCTACGGGAGCTGATCAGCCCAATTCCTGGC 420
Qy 361 GCCACCATGCACACACCTCCAGTGCACCTACGGGAGCTGATCAGCCCAATTCCTGGC 420

Db 421 GAGAACTCGGTCAATCTGACCGAGCCGAGGCTTCCAGAACAAAGGGCTTCACGAATCCC 480
Qy 421 GAGAACTCGGTCAATCTGACCGAGCCGAGGCTTCCAGAACAAAGGGCTTCACGAATCCC 480

Db 481 ATCCAGTTCCTCTCTGTTCTATGGCGGGTACCTTCTCGCTGATGCTGAGGCGCTGG 540
Qy 481 ATCCAGTTCCTCTCTGTTCTATGGCGGGTACCTTCTCGCTGATGCTGAGGCGCTGG 540

Db 541 CATGATACGAACATAGCGCAATGCGGACCAACAGCTCCTCTCATCCAGCGACTCTTG 600
Qy 541 CATGATACGAACATAGCGCAATGCGGACCAACAGCTCCTCTCATCCAGCGACTCTTG 600

Db 601 GTGCAGCAGGTACTGGAGGTGCTCCGAATGGAAGCAACAAAGTCGGAATCGCAGTAC 660
Qy 601 GTGCAGCAGGTACTGGAGGTGCTCCGAATGGAAGCAACAAAGTCGGAATCGCAGTAC 660

Db 661 ACCTGCTGGAGTACGATTTCCGTGTCACCTGCGATCTCAACTACTACGGATCCGGCTGT 720
Qy 661 ACCTGCTGGAGTACGATTTCCGTGTCACCTGCGATCTCAACTACTACGGATCCGGCTGT 720

Db 721 GCCAAGTTCGCGGCGCGGAGGATTCATTGGACACTCGACTTGTCTCGGAGACGGGC 780
Qy 721 GCCAAGTTCGCGGCGCGGAGGATTCATTGGACACTCGACTTGTCTCGGAGACGGGC 780

Db 781 GAAATTAATCTGTTGACCGGATGCGAGCGGATTTACTGTACATACCCAAATGCGCCAAA 840
Qy 781 GAAATTAATCTGTTGACCGGATGCGAGCGGATTTACTGTACATACCCAAATGCGCCAAA 840

Db 841 GGCTGTGAACATGACATTTGGCAAAACCCCAATCAATGCGTTTGCCTGCTGGCTGGAG 900
Qy 841 GGCTGTGAACATGACATTTGGCAAAACCCCAATCAATGCGTTTGCCTGCTGGCTGGAG 900

Db 901 GGAGCCTTGTGCACAGTGGTCTTGGAAACCGAAGTGCATTCATGGCACCTGCACACAA 960
Qy 901 GGAGCCTTGTGCACAGTGGTCTTGGAAACCGAAGTGCATTCATGGCACCTGCACACAA 960

Qy 901 GGAGCCTTGTGCACAGTGGTCTTGGAAACCGAAGTGCATTCATGGCACCTGCACACAA 960

Db 961 CCCTGGACTTGCATCTGCAACGAGGGTTGGGAGGCTTGTACTGCAACCCAGGATCTGAAC 1020

Qy 961 CCCTGGACTTGCATCTGCAACGAGGGTTGGGAGGCTTGTACTGCAACCCAGGATCTGAAC 1020

Db 1021 TACTGCAACCAACACAGCCTGCAAGAATGGCGGAACCTCTTCAACACCCGGGAGGA 1080

Qy 1021 TACTGCAACCAACACAGCCTGCAAGAATGGCGGAACCTCTTCAACACCCGGGAGGA 1080

Db 1081 TTGTACACATCAAAATGCGCTCCAGGATACAGTGGTGATGATGCGAAAATGAGATCTAC 1140

Qy 1081 TTGTACACATCAAAATGCGCTCCAGGATACAGTGGTGATGATGCGAAAATGAGATCTAC 1140

Db 1141 TCCTGCGATGCGGATGTCATCCCTGCCAGAAATGGTGATCTGTCATCGATGAGCCGAC 1200

Qy 1141 TCCTGCGATGCGGATGTCATCCCTGCCAGAAATGGTGATCTGTCATCGATGAGCCGAC 1200

Db 1201 AAAAAACCGGCTCAAGTGTCAATTCGCGCAACGCTGGAGCGAAAGATGTGGAGGAG 1260

Qy 1201 AAAAAACCGGCTCAAGTGTCAATTCGCGCAACGCTGGAGCGAAAGATGTGGAGGAG 1260

Db 1261 AAAGTGCTCAGTGTTCGGACAAACCCCTGTCATCAGGGAATCTGCCGCAACGTTGCTCT 1320

Qy 1261 AAAGTGCTCAGTGTTCGGACAAACCCCTGTCATCAGGGAATCTGCCGCAACGTTGCTCT 1320

Db 1321 GGCTTGGGAACAAAGGTCAGGGCTACAGTGCAGTAATGTCCTTGGCTACAGCGGACCC 1380

Qy 1321 GGCTTGGGAACAAAGGTCAGGGCTACAGTGCAGTAATGTCCTTGGCTACAGCGGACCC 1380

Db 1381 AACTGCGATCTCCAGCTGCAACTGCAAGTCCGAATCCATCAATAACGGTGGAGCTGT 1440

Qy 1381 AACTGCGATCTCCAGCTGCAACTGCAAGTCCGAATCCATCAATAACGGTGGAGCTGT 1440

Db 1441 CAGCGAGCGGAAAGTGTATTGTCAGCGGATTTTCGGGAAACAGATGCGAGACCAAC 1500

Qy 1441 CAGCGAGCGGAAAGTGTATTGTCAGCGGATTTTCGGGAAACAGATGCGAGACCAAC 1500

Db 1501 ATTGACGATGCTTTGGCCACAGTCGAGAACCGAGGACCTGTCATAGATATGCTCAAC 1560

Qy 1501 ATTGACGATGCTTTGGCCACAGTCGAGAACCGAGGACCTGTCATAGATATGCTCAAC 1560

Db 1561 CAATATCGCTGCCAATGCTGCTCCCGTTCCATGCGACCCACTGTAGTACGAAAGTTGAC 1620

Qy 1561 CAATATCGCTGCCAATGCTGCTCCCGTTCCATGCGACCCACTGTAGTACGAAAGTTGAC 1620

Db 1621 TTGTGCTCCTCAGACCTGTCGCAATGGAGGAACCTGCTTGAATCTCAACACGATTAC 1680

Qy 1621 TTGTGCTCCTCAGACCTGTCGCAATGGAGGAACCTGCTTGAATCTCAACACGATTAC 1680

Db 1681 CAGTGCACCTGCTGTCGGGATTTACTGGCAAGGATTCCTCTGTGGACATCGATGAGTGC 1740

Qy 1681 CAGTGCACCTGCTGTCGGGATTTACTGGCAAGGATTCCTCTGTGGACATCGATGAGTGC 1740

Db 1741 AGCAGTGAACCTGTCATAACGGGGGCACTTGCATGAACCGGCTCAATTCGTTGCAATGC 1800

Qy 1741 AGCAGTGAACCTGTCATAACGGGGGCACTTGCATGAACCGGCTCAATTCGTTGCAATGC 1800

Db 1801 GTGTGTGCAATGTTTCAGGGGCAAGCAGTCGATGAGGAGTCCATGATTCGGTGACC 1860

Qy 1801 GTGTGTGCAATGTTTCAGGGGCAAGCAGTCGATGAGGAGTCCATGATTCGGTGACC 1860

Db 1861 TTCATGTCGCCCAATATGAGGCGCACCAACAAAGCGAGAGCGGATGTTTACCAATGCC 1920

Qy 1861 TTCATGTCGCCCAATATGAGGCGCACCAACAAAGCGAGAGCGGATGTTTACCAATGCC 1920

Db 1921 CAGGTAGTCTAAATGCTGTTTCTCCGTTGCGATGCTTGGTGCGGCTTATTGCGGCG 1980

Qy 1921 CAGGTAGTCTAAATGCTGTTTCTCCGTTGCGATGCTTGGTGCGGCTTATTGCGGCG 1980

Db 1981 TGCGTGTCTTCGATGAACGCGAAGCTAAGCTGCTCAGGAAAGACGACGCGAG 2040

Qy 1981 TGCGTGTCTTCGATGAACGCGAAGCTAAGCTGCTCAGGAAAGACGACGCGAG 2040

Db 2041 GCCAGGAGCAGACGACAGACAGATCCGCTGGCCACATGCATCACAATGGCAGGGGTC 2100
| | | | |
Qy 2041 GCCAGGAGCAGACGACAGACAGATCCGCTGGCCACATGCATCACAATGGCAGGGGTC 2100
| | | | |
Db 2101 GGTGAGCTTTGGCTTTCAGCTCTCTGCGGCGCAAAATCGGACAGACAGCGGTCTCAC 2160
| | | | |
Qy 2101 GGTGAGCTTTGGCTTTCAGCTCTCTGCGGCGCAAAATCGGACAGACAGCGGTCTCAC 2160
| | | | |
Db 2161 TTGCATGGCGGCAACCCGATATATCATCAAAAACACCTGGGACAGTCCGTCAACAAT 2220
| | | | |
Qy 2161 TTGCATGGCGGCAACCCGATATATCATCAAAAACACCTGGGACAGTCCGTCAACAAT 2220
| | | | |
Db 2221 TGTGCTTTCAGACAGACAGCGCGCGCGCGGCGGACAGCGGCGGAGAGTCTCATGAC 2280
| | | | |
Qy 2221 TGTGCTTTCAGACAGACAGCGCGCGCGCGGCGGACAGCGGCGGAGAGTCTCATGAC 2280
| | | | |
Db 2281 GGGGATATGTCGCTCGTGGCGGATACAAACATGGCACTGAGCTTTTGTGGCT 2340
| | | | |
Qy 2281 GGGGATATGTCGCTCGTGGCGGATACAAACATGGCACTGAGCTTTTGTGGCT 2340
| | | | |
Db 2341 CCGCTACAAAGAGCCAACTCCAAAGCACTCAACACCGATCCAGCTCATGACCGC 2400
| | | | |
Qy 2341 CCGCTACAAAGAGCCAACTCCAAAGCACTCAACACCGATCCAGCTCATGACCGC 2400
| | | | |
Db 2401 GGTTCGCGCGGAGCAGCTCAGCCAAAGGAGCGTCTGGCGGAGACCGGAGCGCGGAG 2460
| | | | |
Qy 2401 GGTTCGCGCGGAGCAGCTCAGCCAAAGGAGCGTCTGGCGGAGACCGGAGCGCGGAG 2460
| | | | |
Db 2461 GGCAGAGAGATCTCTGTTTGAAGGAGGTTCTCTAGCAGCAGTGGCCCTCGATTG 2520
| | | | |
Qy 2461 GGCAGAGAGATCTCTGTTTGAAGGAGGTTCTCTAGCAGCAGTGGCCCTCGATTG 2520
| | | | |
Db 2521 GCGGCGGCGGAGTGGCGGAGCGCTGTTCATCCAGCTAATGGCTGACAGCTTCGCGACG 2580
| | | | |
Qy 2521 GCGGCGGCGGAGTGGCGGAGCGCTGTTCATCCAGCTAATGGCTGACAGCTTCGCGACG 2580
| | | | |
Db 2581 GGCAGCGGAGCGGAGCGGCGGACAGCAGCAGTCCGTGCTGGCGACCTCCGATATG 2640
| | | | |
Qy 2581 GGCAGCGGAGCGGAGCGGCGGACAGCAGCAGTCCGTGCTGGCGACCTCCGATATG 2640
| | | | |
Db 2641 TAATCCAAAATCCGGAAGGCTCTGTGAATCCGGAAGATCCGATGGAGAGGTCG 2700
| | | | |
Qy 2641 TAATCCAAAATCCGGAAGGCTCTGTGAATCCGGAAGATCCGATGGAGAGGTCG 2700
| | | | |
Db 2701 ACACACATACACAAGAAAAGTGGGTTCAAAATGTAGAGAGAGCGCCAAAT 2760
| | | | |
Qy 2701 ACACACATACACAAGAAAAGTGGGTTCAAAATGTAGAGAGAGCGCCAAAT 2760
| | | | |
Db 2761 GTTGTGTGATGTAAGCAGTTAGTGTACGAAAAATGAAAAATCTGTAAAGGACATA 2820
| | | | |
Qy 2761 GTTGTGTGATGTAAGCAGTTAGTGTACGAAAAATGAAAAATCTGTAAAGGACATA 2820
| | | | |
Db 2821 ACTGTAAACCTCCATAAAATTTGTATATGAAATTAAGCAAAAGCTGTGACCGCTTTC 2880
| | | | |
Qy 2821 ACTGTAAACCTCCATAAAATTTGTATATGAAATTAAGCAAAAGCTGTGACCGCTTTC 2880
| | | | |
Db 2881 GATCCGAATTC 2892
| | | | |
Qy 2881 GATCCGAATTC 2892
| | | | |

CC TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
CC NUMBER OF SEQUENCES: 52
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Foley & Lardner
CC STREET: 1800 Diagonal Road, Suite 500
CC CITY: Alexandria
CC STATE: VA
CC COUNTRY: USA
CC ZIP: 22313-0299
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/232,463
CC FILING DATE:
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US/07/935,313
CC FILING DATE:
CC APPLICATION NUMBER: EP 91 114 300, 6
CC FILING DATE: 26-AUG-1991
CC ATTORNEY/AGENT INFORMATION:
CC NAME: BENT, Stephen A.
CC REGISTRATION NUMBER: 29,768
CC REFERENCE/DOCKET NUMBER: 30472/114 IMMU
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (703)836-9300
CC TELEFAX: (703)683-4109
CC TELEX: 899149
CC INFORMATION FOR SEQ ID NO: 14:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 7218 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC IMMEDIATE SOURCE:
CC CLONE: pTZsp1-F15
CC Sequence 7218 bp; 1944 A; 1491 C; 1486 G; 1929 T; 368 other;
S0

Query Match 2.5%; Score 71; DB 1; Length 7218;
Best Local Similarity 3.1%; Pred. No. 6,666-32;
Matches 12; Conservative 215; Mismatches 136; Indels 0; Gaps 0;

Db 1053 AGGAGCTTGGAGATTT 1112
| | | | |
Cp 2835 AGGAGCTTACAGATATGCGCTGTACAGATTTTCATTTTCGTAGAGACTAACTGCT 2776
| | | | |
Db 1113 TCAATCAACAACATTTTGGCGCTCTCACAATTTGAACCAACCAAGCTTTTCT 2716
| | | | |
Cp 2775 TCAATCAACAACATTTTGGCGCTCTCACAATTTGAACCAACCAAGCTTTTCT 2716
| | | | |
Db 1113 TTTGTATGTCGTCACTCTCCATCGGATTTCTCCGATTTTACAGAGAGCCCTTC 2656
| | | | |
Cp 2715 TTTGTATGTCGTCACTCTCCATCGGATTTCTCCGATTTTACAGAGAGCCCTTC 2656
| | | | |
Db 1233 TTTGTATGTCGTCACTCTCCATCGGATTTCTCCGATTTTACAGAGAGCCCTTC 1292
| | | | |
Cp 2655 GATTTTGGAGTTACATATGCGAGTGCAGCAGCAGATCTGCTGTGGCGCT 2596
| | | | |
Db 1293 TTTGTATGTCGTCACTCTCCATCGGATTTCTCCGATTTTACAGAGAGCCCTTC 1352
| | | | |
Cp 2595 CCGCGCTCCGCGTCCGCTGCGAGAGCTGACAGCATTTAGTGGATGAAAGAGCTCCGCG 2536
| | | | |
Db 1353 TTTGTATGTCGTCACTCTCCATCGGATTTCTCCGATTTTACAGAGAGCCCTTC 1412
| | | | |
Cp 2535 CACTCCCGCGCGCAGCAGAGGCGCAAGCTGCTACAGTAGAGAACCTCGCTAAAC 2476
| | | | |
Db 1413 TTTGTATGTCGTCACTCTCCATCGGATTTCTCCGATTTTACAGAGAGCCCTTC 1435
| | | | |
Cp 2475 AGAGATCTCTGCGCTCCGCGC 2453
| | | | |

Best Local Similarity	14.8%;	Pred. No. 4,966-12;
Matches	31; Conservative	87; Mismatches 88; Indels 3; Gaps 22;

Db 6 MSSSVYRRATSCNDKRAKKDNTSSWPTDCCNRTWCGDPTLYRYANNDSGHKXYSAN 65
 Cp 498 CGAGAGGGGGAACCTAGTAGGGGATTCGTGAMACCCTTGTCTGAGAGCGTGGCGCTGCT 4398
 Db 66 YN - YGCNNVGAANTHYTHTNVSGADSKTYTDS - YNAGTSSNSGDTGNSRGADSYGS 122
 Cp 438 CAGATTGACGAGATTCTCGCGAGAAATGGCGGTATACAGTCCCGCGTAGGTGCACTGGGA 379
 Db 123 SKTAVTSNNPFGKTAANNAVDSRNMGDSVSGDKTKKHAANSADGKYGSKNNGDRNNRYG 182
 Cp 378 GGTGCTGTCGATGCTGGCGCTGATGCTTTAGCGCAGACGCGGAACCGCTTGCAGCT 319
 Db 183 TGTSAVSNNGCGNKRDPVSYAANNKCG 211
 Cp 318 GCCCAGGCACTTGGCCGCTGCTCCGCTCCG 290

RESULT	5
ID	US-08-597-545-4 STANDARD; DNA; UNC; 1573 BP.

DT 01-JAN-1900
DE Sequence 4, Application US/08597545
CC Sequence 4, Application US/08597545
CC Patent No. 5580738
CC GENERAL INFORMATION:

CC GENERAL INFORMATION:
CC APPLICANT: LABORDA, Jorge
CC TITLE OF INVENTION: Delta-like Gene Expressed In
CC TITLE OF INVENTION: Neuroendocrine Tumors
CC NUMBER OF SEQUENCES: 9
CC CORRESPONDENCE ADDRESS:

CC ADDRESSEE: Foley & Lardner
CC STREET: 3000 K Street, N.W., Suite 5000
CC CITY: Washington, D.C.

```
CC      COMPUTER READABLE FORM:
CC      MEDIUM TYPE: Floppy disk
CC      COMPUTER: IBM PC compatible
CC      OPERATING SYSTEM: PC-DOS/MS-DOS
CC      SOFTWARE: PatentIn Release #1.0, Version #1.25
CC      CURRENT APPLICATION DATA:
CC      APPLICATION NUMBER: US/08/597,545
```

CC FILING DATE: 435
CC CLASSIFICATION:
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/989,537
CC FILING DATE: 11-DEC-1992
CC ATTORNEY/AGENT INFORMATION:

CC NAME: BENT, Stephen A.
CC REGISTRATION NUMBER: 29,768
CC REFERENCE/DOCKET NUMBER: 40399/166 NIND
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (202)672-5300
CC TELEFAX: (202)672-5399

```
CC INFORMATION FOR SEQ ID NO: 4:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1573 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: double
CC TOPOLOGY: linear
CC ORIGINAL SOURCE:
```

CC STRAIN: Mouse D1k
SQ Sequence 1573 BP; 327 A; 507 C; 419 G; 320 T; 0 other;

Query Match	1.5%;	Score 42;	DB 1;	Length 1573;
Best Local Similarity	58.3%;	Pred. No. 4.96e-12;		
Matches 200; Conservative	0;	Mismatches 140;	Indels 3;	Gaps 2;

Db 234 ATGATTTCTGCGAGGCTGACAAATGTTCTGCAGGTGCCATGTTGGCTGGAGGTTCCCTCT 293

851 ATGCACATTGCGACAACCCATCATCGCTTGCCCACTGGCGCTGGAAGGAGCCCTGT 910

Db 294 GTGACAGTGTAACTGCCCTGGCTGTGTCAATGAGTGTGCAAGSAACCATGGCAGT 3533

911 GCACGAGTGGCTCTGTGAACCGAACGTCATCCATGGCACCCTGCACCAAAACCTTGGACTT 910

971 GCATCTGCCACGAGGGTTGGGGAGGCGTTTACTGCAACCAAGATCTGAACCTACTGCAACCA 103

Db 414 --CA-ACCCCTCCGCCACATGGAAGTGGCTGGACCTGGAGAAAGCCAGTACGAAT 470

1031 ACCACGACCCCTSCAAGAAIGCCGGACCTGCTTCACACCGCGGAGGATTTGTACAT 1039

1091 GCGAAATGCGCTCCAGGATACAGTGTGATGATTCGAAATAGAGATCTACTCTCGGATG 115

Db 531 TCATGGTTCTCCCTGCCAGCAGGAGGCGCCTCGGTGGANGA 573

1151 CCGAGTCAATCCCTGCCAGATGTGGTACCTGCATGATGA 1193

RESULT	6
ID	US-08-457-135-4
STANDARD:	DNA
TTC:	1573
BP:	

AC XXXXXX
DT 01-JAN-1900

RESULT	6
ID	US-08-457-135-4 STANDARD; DNA; UNC; 1573 BP.
AC	xxxxxx
DT	01-JAN-1900
DE	Sequence 4, Application US/08457135.
CC	Sequence 4, Application US/08457135
CC	Patent No. 5644031
CC	GENERAL INFORMATION:

CC GENERAL INFORMATION:
CC APPLICANT: LABORDA, Jorge
CC TITLE OF INVENTION: Delta-Like Gene Expressed In
CC TITLE OF INVENTION: Neuroendocrine Tumors
CC NUMBER OF SEQUENCES: 9
CC CORRESPONDENCE ADDRESS:
CC Address: P.O. Box 11460
CC

CC ADDRESSEE: Foley & Lardner
CC STREET: 3000 K Street, N.W., Suite 500
CC CITY: Washington, D.C.
CC COUNTRY: USA
CC ZIP: 20007-5109
CC COMMENTER: DEANARIE ENOM.

```
CC      COMPUTER READABLE FORM:
CC      MEDIUM TYPE: Floppy disk
CC      COMPUTER: IBM PC compatible
CC      OPERATING SYSTEM: PC-DOS/MS-DOS
CC      SOFTWARE: PatentIn Release #1.0, Version #1.25
CC      CURRENT APPLICATION DATA:
```

CC APPLICATION NUMBER: US/08/457,135
CC FILING DATE: 01-JUN-1995
CC
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 07/989,537
CC FILING DATE: 11-DEC-1992
CC
CC ATTORNEY/AGENT INFORMATION:
CC

CC NAME: BENT, Stephen A.
CC REGISTRATION NUMBER: 29,768
CC REFERENCE/DOCKET NUMBER: 40399/304/NHIE
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (202)672-5300
CC TELEFAX: (202)672-5399

```
CC      TELEX: 904136
CC      INFORMATION FOR SEQ ID NO: 4
CC      SEQUENCE CHARACTERISTICS:
CC      LENGTH: 1573 base pairs
CC      TYPE: nucleic acid
CC      STRANDEDNESS: double
CC      TOPOLOGY: linear
CC      ORIGINAL SOURCE:
```

CC STRAIN: Mouse D1k
SQ Sequence 1573 BP; 327 A; 507 C; 419 G; 320 T; 0 other;

Query Match	1.5%;	Score 42;	DB 1;	Length 1573;
Best Local Similarity	58.3%;	Pred. No. 4.96e-12;		


```
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 7218 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC IMMEDIATE SOURCE:
CC CLONE: pT29pt-F15
SQ Sequence 7218 BP; 1944 A; 1491 C; 1486 G; 1929 T; 368 other;

Query Match 1.38; Score 38; DB 1; Length 7218;
Best Local Similarity 2.08; Pred. No. 1.74e-09;
Matches 4; Conservative 116; Mismatches 82; Indels 0; Gaps 0;

Db 1062 GCGATYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1121
||| |: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 330 GCGGTTTCGCTGCTAAAGCACTACCAAGCCACCATCGACACCACTCCCAAGTGAC 389
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1122 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1181
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 390 CTACGGGACGTGATCAGCCCACTTCGCGGAGAACTCGGTCAATCGACGAGCCCA 449
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1182 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1241
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 450 GCGCTTCAGAACAGGCTTCACGAATCCATCCAGTTCCTCTCGTTCATGGCC 509
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1242 YYYYYYYYYYYYYYYYYYYYYY 1263
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 510 GGGTACCTTCGCTGATCGTC 531

RESULT 11
ID PCT-US96-02331-14 STANDARD; DNA; UNC; 4255 BP.
AC xxxxxx
DT 01-JAN-1900
DE Sequence 14, Application PC/TUS9602331.
CC Sequence 14, Application PC/TUS9602331
CC GENERAL INFORMATION:
CC APPLICANT: The Board of Trustees of the Leland Stanford Junior
CC APPLICANT: University
CC APPLICANT: Board of Regents, The University of Texas System
CC TITLE OF INVENTION: Methods and Compositions for Altering
CC TITLE OF INVENTION: Sexual Behavior
CC NUMBER OF SEQUENCES: 15
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Dehlinger & Associates
CC STREET: 350 Cambridge Avenue, Suite 250
CC CITY: Palo Alto
CC STATE: CA
CC COUNTRY: USA
CC ZIP: 94306
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US96/02331
CC FILING DATE: 09-FEB-1996
CC CLASSIFICATION:
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/386,495
CC FILING DATE: 10-FEB-1995
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Sholtz, Charles K.
CC REGISTRATION NUMBER: 38,615
CC REFERENCE/DOCKET NUMBER: 8600-0153.41
CC TELEPHONE: (415) 324-0880
CC TELEFAX: (415) 324-0960
CC INFORMATION FOR SEQ ID NO: 9:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 4255 base pairs
CC STRANDEDNESS: double
CC TYPE: nucleic acid
CC TOPOLOGY: unknown
CC MOLECULE TYPE: CDNA to mRNA
CC HYPOTHETICAL: NO
CC ANTI-SENSE: NO
CC ORIGINAL SOURCE:
CC INDIVIDUAL ISOLATE: Fru#1 CDNA
CC FEATURE:
CC NAME/KEY: CDS

CC STRANDEDNESS: double
CC TOPOLOGY: unknown
CC MOLECULE TYPE: CDNA to mRNA
CC HYPOTHETICAL: NO
CC ANTI-SENSE: NO
CC ORIGINAL SOURCE:
CC INDIVIDUAL ISOLATE: fruitless transcript in Fig. 7E
CC FEATURE:
CC NAME/KEY: CDS

CC LOCATION: 1507..4032
SQ Sequence 4255 BP; 1273 A; 1130 C; 1000 G; 852 T; 0 other;

Query Match 1.18; Score 32; DB 2; Length 4255;
Best Local Similarity 92.18; Pred. No. 8.10e-06;
Matches 35; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Db 2073 CTCAGCGCAGCAGCGCGCGCGCGCAGCAGCAGCGG 2110
||||| |: : : : : : : : : : : : : : : : : : : : : : : : :
QY 2226 CTCAGCAGCAGCAGCGCGCGCGCGCAGCAGCGG 2263

RESULT 12
ID PCT-US96-02331-9 STANDARD; DNA; UNC; 4835 BP.
AC xxxxxx
DT 01-JAN-1900
DE Sequence 9, Application PC/TUS9602331.
CC Sequence 9, Application PC/TUS9602331
CC GENERAL INFORMATION:
CC APPLICANT: The Board of Trustees of the Leland Stanford Junior
CC APPLICANT: University
CC APPLICANT: Board of Regents, The University of Texas System
CC TITLE OF INVENTION: Methods and Compositions for Altering
CC TITLE OF INVENTION: Sexual Behavior
CC NUMBER OF SEQUENCES: 15
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Dehlinger & Associates
CC STREET: 350 Cambridge Avenue, Suite 250
CC CITY: Palo Alto
CC STATE: CA
CC COUNTRY: USA
CC ZIP: 94306
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US96/02331
CC FILING DATE: 09-FEB-1996
CC CLASSIFICATION:
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/386,495
CC FILING DATE: 10-FEB-1995
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Sholtz, Charles K.
CC REGISTRATION NUMBER: 38,615
CC REFERENCE/DOCKET NUMBER: 8600-0153.41
CC TELEPHONE: (415) 324-0880
CC TELEFAX: (415) 324-0960
CC INFORMATION FOR SEQ ID NO: 9:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 4835 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: double
CC MOLECULE TYPE: CDNA to mRNA
CC HYPOTHETICAL: NO
CC ANTI-SENSE: NO
CC ORIGINAL SOURCE:
CC INDIVIDUAL ISOLATE: Fru#1 CDNA
CC FEATURE:
CC NAME/KEY: CDS
```


CC Sequence 1, Application US/08385500
CC Patent No. 5712117
CC GENERAL INFORMATION:
CC APPLICANT: Sprecher, Cindy A.
CC TITLE OF INVENTION: CYTOPLASMIC ANTIPROTEINASE-2 AND
CC TITLE OF INVENTION: CYTOPLASMIC ANTIPROTEINASE-3 AND CODING SEQUENCES
CC NUMBER OF SEQUENCES: 16
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Townsend and Townsend Kourie and Crew
CC STREET: Steuart Street Tower, One Market Plaza
CC CITY: San Francisco
CC STATE: California
CC COUNTRY: US
CC ZIP: 94105-1493
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/385,500
CC FILING DATE:
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Parmelee, Steven W.
CC REGISTRATION NUMBER: 31,990
CC REFERENCE/DOCKET NUMBER: 13952-21
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (206) 467-9600
CC TELEFAX: (415) 543-5043
CC INFORMATION FOR SEQ ID NO: 1:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1425 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: cDNA
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 92..1213
CC OTHER INFORMATION: /product= "CYTOPLASMIC
CC OTHER INFORMATION: ANTIPROTEINASE-2 PROTEIN"
CC OTHER INFORMATION: ANTIPROTEINASE-2 PROTEIN"
SQ Sequence 1425 BP; 437 A; 301 C; 357 G; 330 T; 0 other;

Query Match 1.0%; Score 28; DB 1; Length 1425;
Best Local Similarity 91.2%; Pred. No. 1.67e-03;
Matches 31; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Db 37 CAGCAGCGCGCGCGCGCGCGCGCAGCAGC 70
QY 2228 CAGCAGCGCGCGCGCGCGCGCGCAGCAGC 2261
||||||| || ||||||||||||||||||| ||

Search completed: Sat Jun 13 10:10:43 1998
Job time : 201 secs.

 (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

Mparch_n n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Sat Jun 13 08:37:40 1998; Maspar time 2790.42 Seconds
 1384.359 Million cell updates/sec

Tabular output not generated.

Title: >US-08-083-590A-1
 Description: (1-2892) from US08083590A.seq
 Perfect Score: 2892
 N.A. Sequence: 1 GAATTCGGAGGATTTATCA.....GCCGTTTCGATCCGGAATTC 2892
 Comp: CTTAAGCCTCTTATATAGT.....CGGCAAGCTTAGGCTTAG

Scoring table: TABLE default
 Gap 6

Mmatch STD : Dbase 0; Query 0

Searched: 1759237 seqs, 667866413 bases x 2

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database:

emb1-est54
 1:em_est1 2:em_est3

Database:

genbank-est106
 3:gb_est1 4:gb_est10 5:gb_est11 6:gb_est12 7:gb_est13
 8:gb_est14 9:gb_est15 10:gb_est16 11:gb_est17
 12:gb_est18 13:gb_est19 14:gb_est2 15:gb_est20
 16:gb_est21 17:gb_est22 18:gb_est3 19:gb_est4 20:gb_est5
 21:gb_est6 22:gb_est7 23:gb_est8 24:gb_est9 25:gb_est9s
 26:gb_ests

Statistics: Mean 11.911; Variance 2.337; scale 5.097

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	436	15.1	442	13	AA538738	LD18334.5prtime LD Dros 0.00e+00
2	377	13.0	379	13	AA540186	LD19584.5prtime LD Dros 0.00e+00
3	150	5.2	568	12	AA142228	CK00091.3prtime CK Dros 2.05e-235
4	78	2.7	567	12	AA015264	mh30d12.rl Soares mous 5.84e-93
5	60	2.1	252	13	AA754459	97SN1787 Rice Immature 6.76e-60
6	57	2.0	247	13	AA754458	97SN1784 Rice Immature 1.40e-54
7	59	2.0	252	13	AA754459	97SN1787 Rice Immature 4.07e-58
8	55	1.9	570	21	AA051023	mg73f08.rl Soares mous 4.48e-51
9	51	1.8	247	13	AA054458	97SN1784 Rice Immature 3.60e-44
10	51	1.8	602	23	AA046860	zf14hl1.rl Soares feta 1.82e-42
11	50	1.7	445	10	AA619107	yf66f10.rl Homo sapien 3.60e-42
12	42	1.5	445	10	AA619107	vo68b02.rl Soares mous 2.96e-29
13	42	1.5	485	23	AA024229	mh99e03.rl Soares mous 2.96e-29

14	37	1.3	589	24	AA238559	my35e01.rl Barstead mo 1.89e-21
15	39	1.3	592	23	AA153148	mr47c02.rl Stratiogene 1.60e-24
16	39	1.3	642	21	AA048742	mj44C04.rl Soares mous 1.60e-24
17	39	1.3	766	23	AA165945	ms50g09.rl Life Tech m 1.60e-24
18	39	1.3	2275	12	AF034173	Homo sapiens ntcon2 co 1.60e-24
19	36	1.2	378	12	AA141415	CK01709.5prtime CK Dros 6.08e-20
20	36	1.2	390	10	W13561	ma85d08.rl Soares mous 5.51e-17
21	34	1.2	527	23	AA171296	ms57a11.rl Life Tech m 1.87e-18
22	35	1.2	786	23	AA171309	ms57e11.rl Life Tech m 1.87e-18
23	35	1.2	2275	12	AF034173	Homo sapiens ntcon2 co 1.87e-18
24	32	1.1	254	14	R50365	yj59e07.s1 Homo sapien 4.08e-14
25	32	1.1	269	22	AA059314	zf66b05.s1 Soares ret1 4.08e-14
26	33	1.1	328	18	H39899	yf02h01.s1 Homo sapien 1.54e-15
27	31	1.1	391	11	AA683105	ae81h06.s1 Stratiogene 1.02e-12
28	31	1.1	416	14	R72432	yj9e005.s1 Homo sapien 4.08e-14
29	31	1.1	422	14	R50026	yj54a01.s1 Homo sapien 4.08e-14
30	32	1.1	459	14	RA6751	zr98c02.s1 NC1_CGAP_GC 1.02e-12
31	31	1.1	470	10	AA215625	LD05485.5prtime LD Dros 1.02e-12
32	31	1.1	518	12	AA246536	342820.TPB CTR978SKA1 4.08e-14
33	32	1.1	607	25	B16636	Human DRES 30 mRNA seq 1.54e-15
34	33	1.1	729	12	HSU69180	mqj0e06.rl Soares 2NDM 2.40e-11
35	30	1.0	136	24	AA139263	md38g09.rl Soares feta 5.30e-10
36	29	1.0	297	22	W67931	mh49e03.s1 Soares mous 2.40e-11
37	30	1.0	337	23	AA020103	yx75c06.s1 Homo sapien 2.40e-11
38	30	1.0	346	19	N24541	H. sapiens (D11S906) D 2.40e-11
39	30	1.0	353	26	HS107XC7	y99d07.rl Homo sapien 2.40e-11
40	30	1.0	390	3	T49928	y128h01.rl Homo sapien 5.30e-10
41	29	1.0	414	18	H16156	zg79b12.s1 Soares feta 5.30e-10
42	29	1.0	421	12	AA733056	zb06h12.rl Soares feta 5.30e-10
43	29	1.0	444	20	W16654	yn65b09.rl Homo sapien 5.30e-10
44	29	1.0	457	18	H20919	yo63b10.rl Homo sapien 5.30e-10
45	29	1.0	487	18	H42553	

ALIGNMENTS

RESULT 1
 LOCUS AA538738 442 bp mRNA EST 19-JAN-1998
 DEFINITION LD18334.5prtime LD Drosophila melanogaster embryo Bluescript
 Drosophila melanogaster cDNA clone LD18334 5prtime similar to
 X05140: Drosophila melanogaster mRNA fragment of DL locus with
 EGF-like repeats (EGF-epidermal growth factor); Y00222:
 D.melanogaster DL mRNA for delta protein, involved in neurogenesis,
 mRNA sequence.

ACCESSION AA538738
 NID 92285254
 KEYWORDS EST.
 SOURCE fruit fly.
 ORGANISM Drosophila melanogaster

Eukaryote; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 Pterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea;
 Drosophilidae; Drosophila.

REFERENCE 1 (bases 1 to 442)

Harvey,D., Hong,L., Evans-Holm,M., Pendleton,J., Su,C.,
 Broksstein,P., Lewis,S. and Rubin,G.M.

BDGP/HMT Drosophila EST Project

Unpublished (1997)

TITLE JOURNAL

COMMENT

Contact: Harvey, D.
 G. M. Rubin-Molecular and Cell Biology
 University of California Berkeley
 539 USA, Berkeley, CA 94720-3200, USA
 Fax: 510 643 9947
 Email: <http://fruitfly.berkeley.edu/EST>, estfruitfly.berkeley.edu
 Plate: 183 row C column: 10
 High quality sequence stop: 390.
 Location/Qualifiers
 1..442
 /organism="Drosophila melanogaster"
 /note="Organ: embryo; Vector: Bluescript SK; Site: 1;
 EcoRI, Site2: XhoI; Constructed using Strategene ZAP-cDNA
 Synthesis Kit; Oligo dt-primed and directionally cloned at
 EcoRI and XhoI in Bluescript SK(+/-)"

FEATURES
 source

/cultivar="MilYang23"
/note="Vector: pBluescript SK(+); Site 1: EcoRI; Site 2: XhoI; Directional cDNA library inserted into lambda ZAPII vector at 5' end with EcoRI and 3' end with Xho I site."
/db_xref="taxon:4530"
/clone="97SN1787"
/clone_lib="Rice Immature Seed Lambda ZAPII cDNA Library"
/tissue_type="Immature Seed"
/dev_stage="5 days after pollination"
/lab_host="E. coli SOLR"
BASE COUNT 5 a 21 c 12 g 35 t 179 others
ORIGIN

Query Match 2.08; Score 59; DB 13; Length 252;
Best local Similarity 11.78; Pred. No. 4,07e-58;
Matches 25; Conservative 11; Mismatches 75; Indels 3; Gaps 3;

Db 26 HGNBWWCVASHGNYMSVHNCBTRGTHDCKNVNMSTWGTVMNBVMBVBYBNT 85
QY 1484 CGAGATGCCGACCATGACGATGCTTGGCCACGACGAGACGAGGACCT 1543
DB 86 KVDVGNHTRCSRWBYTRMAHYHDYTCBBYNNNDYHMBBYBYTCMTCTMGCMBHY 145
QY 1544 GCATAGATATGCT-CAACCAATATCGCTGCCAATGCGTCCCGGTTCATGACGACCCAC 1602
DB 146 NTKC-TASGHTSTNDYKSTNTWGTBSYDKSMHYNCSBVKYHTKYSTTATRSYT 204
QY 1603 TGTAGTAGCAAAAGTTGCTGCTCATC-AGACCGTGTGCCAATGAGAGAACCTGCT 1661
DB 205 CVRYCVMNMTTKYKKYHYVBBGCHBTDSKCT 238
QY 1662 GAATCTCAACAGATTACAGTGCACCTGCTCT 1695

RESULT 8
LOCUS AA051023 570 bp mRNA EST 09-SEP-1996
DEFINITION mg73f08.r1 Soares mouse embryo NbMEL3.5 Mus musculus cDNA.
clone 438663 5' similar to gb:X80903 M.musculus mRNA for Delta-like 1 protein (MOUSE); mRNA sequence.
ACCESSION AA051023
NID 91530694
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 570)
AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisel, S., Kucaba, R., Lacy, M., Le, M., Martin, J., Morris, M., Scheinberg, R., Steptoe, M., Tan, F., Underwood, R., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.
TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Maria M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through INL; contact the IMAGE Consortium (info@image.lnl.gov) for further information.
MGI:263999
Seq primer: -28M13 rev2 from Amersham
High quality sequence stop: 287.
Location/Qualifiers
1..570
/organism="Mus musculus"
/strain="C57BL/6J"
/note="Vector: pT73D-Pac (Pharmacia) with a modified

polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTTACCATCTCAAGTGGAGGCGCCGCGAAATTTTTTTTTTTTTTTT T 3'], on equal amounts of mRNA from 2 13.5dpc and 2 14.5dpc embryos [total RNA provided by Minoru Ko, Wayne State Univ., from 2 1]; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo.
/db_xref="taxon:10090"
/clone="438663"
/clone_lib="Soares mouse embryo NbMEL3.5 14.5"
/sex="unknown"
/tissue_type="embryo"
/dev_stage="13.5-14.5dpc total fetus"
/lab_host="DH10B"
BASE COUNT 112 a 180 c 156 g 122 t
ORIGIN

Query Match 1.98; Score 55; DB 21; Length 570;
Best local Similarity 66.58; Pred. No. 4.48e-51;
Matches 151; Conservative 0; Mismatches 72; Indels 4; Gaps 3;

Db 283 GCTTTGGCCTCGAGACCTCTTTGGCGTATGCTTACGACTACGAGCGCTGT 342
QY 311 GCTTGGGCGAGTGCACAGCGGCTTTCGCTCTCAAGACATACGAGCGCACATCG 370
DB 343 CACGGGAGACCACTGACGACGAGGAGTGTGTACAGCCAGCTCGGTGTGACACT 402
QY 371 ACACCACCTCCAGTACGACGAGGAGGAGTGTGTGACGACCCATCTCGGAGACTGG 430
DB 403 TCAGCTGCTGTGATGCC-GCAGATC--GAACCGGC-CTTCAGCAACCCATCCGATTC 458
QY 431 TCATCTGACCGAGCGCCAGCGCTTCAGACAGGCGTTACGATCCATCCAGTTCC 490
DB 459 CTTTGGCTTACCTGCGCAGGTATCTTCTCTGATCATTTGAAGCC 505
QY 491 CTTTCTGTTCTCATGGCGGCTACTTCTCTGATCTGAGGCC 537

RESULT 9
LOCUS AA754458 247 bp mRNA EST 20-JAN-1998
DEFINITION 97SN1784 Rice Immature Seed Lambda ZAPII cDNA Library Oryza sativa cDNA clone 97SN1784, mRNA sequence.
ACCESSION AA754458
NID 92801164
KEYWORDS EST.
SOURCE rice.
ORGANISM Oryza sativa
Eukaryotes; Viridiplantae; Charophyta/Embryophyta group; Embryophyta; Tracheophyta; seed plants; Magnoliophyta; Liliopsida; Poales; Poaceae; Oryza.

REFERENCE 1 (bases 1 to 247)
AUTHORS Nahm, B.H., Kim, J.K., Cheong, J.J., Kim, S.I., Hahn, T.R., Moon, E.P., Kim, W.T., Kim, W.Y., Yang, M.S., Park, R.D., Sohn, D.I., Kang, K.Y., Lee, M.C. and Eun, M.Y.
TITLE Large-scale Sequencing Analysis of ESTs from Rice Immature Seed
JOURNAL Unpublished (1998)
COMMENT Contact: Eun M.Y.
Department of Cyto genetics
National Inst. of Agri. Sci. and Tech, RDA
Suwon, Kyunggi-do, Korea
Tel: 82 331 290 0301
Fax: 82 331 290 0307
Email: myeun@sun20.asstl.re.kr
Submitted by Baek Hie Nahm, Dept of Biological Science, Myongji University, Yongin, Korea. 449-728 bina@mbioserver.myongji.ac.kr
Seq primer: M13 Reverse Primer.
Location/Qualifiers

Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 850L, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu

This clone is available royalty-free through LLNL; contact the
IMAE Consortium (info@image.llnl.gov) for further information.
GG1:586595

Contact: Maria M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:275988
Possible reversed clone: similarity on wrong strand

Location/Qualifiers

WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
MGI:369898

Seq primer: -28m13 rev1 ET from Amersham
High quality sequence stop: 381.

FEATURES

source

1..592

location/Qualifiers

/organism="Mus musculus"

/strain="NIH/Swiss"

/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
XhoI; Cloned unidirectionally. Primer: Oligo dt. 93
pooled NIH/Swiss 13 day embryo hearts. Average insert
size: 1.0 kb; Uni-ZAP XR Vector: -5' adaptor sequence: 5'
GAATCGGCACGAG 3' -3' adaptor sequence: 5'
CTCGAGTTTCTTTTCTTTT 3'"

/db_xref="taxon:10090"

/clone="604466"

/clone_lib="Stratagene mouse heart (#937316)"

/sex="pooled"

/dev_stage="13 day embryos"

/lab_host="SOLR (kanamycin resistant)"

<1..>592

ORIGIN

BASE COUNT

106 a 196 c 176 g 114 t

Query Match

Best local Similarity 63.8%; Score 39; DB 23; Length 592;

Matches 90; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

DB

239 ATGGATTCTGCGAGGCTGACATGCTGCGAGGTCATGTTGGCTGGAGGTCCTCT 298

OY

851 ATGGACATTGCGACAAACCAATCAATGCTTTGCCACTGGCTGGAGGAGCCCTGT 910

DB

299 GTGACAAGTGTGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 358

OY

911 GCAAGAGTGGCTGTGTAAGCACTGATGATGATGATGATGATGATGATGATGATG 970

DB

359 GCATCTGCAAGATGGCTGGG 379

OY

971 GCATCTGCAAGAGGCTGGG 991

Search completed: Sat Jun 13 09:56:28 1998

Job time : 4728 secs.

118184969 CA: 118(19)184969g JOURNAL
 Mouse mammary tumor gene int-3: a member of the notch gene family
 transforms mammary epithelial cells
 AUTHOR(S): Robbins, Joan; Blondel, Bruno J.; Gallahan, Daniel; Callahan, Robert
 LOCATION: Lab. Tumor Immunol. Biol., Natl. Cancer Inst., Bethesda, MD, 20892, USA
 JOURNAL: J. Virol. DATE: 1992 VOLUME: 66 NUMBER: 4 PAGES: 2594-9
 CODEN: JOVIAM ISSN: 0022-538X LANGUAGE: English
 SECTION:
 CA203004 Biochemical Genetics
 CA213XXX Mammalian Biochemistry
 IDENTIFIERS: mouse mammary tumor gene int3 transformation, epithelium transformation mouse notch like gene
 DESCRIPTORS:
 Epithelium...
 cell, of mammary gland of mouse, gene int-3 transformation of
 Deoxyribonucleic acid sequences...
 for gene int-3 protein of mouse mammary tumor
 Proteins, specific or class, gene int-3...
 gene for, of mouse mammary tumor, sequence of and epithelial cell transformation by
 Virus, animal, murine mammary tumor...
 gene int-3 integration site for, sequence of and epithelial cell transformation by
 Mouse...
 gene int-3 of mammary tumor of, sequence of and epithelial cell transformation by
 Gene, animal, Notch...
 mouse mammary tumor gene int-3 like, of Drosophila melanogaster, sequence of and epithelial cell transformation by
 Mammary gland, neoplasm...
 notch family gene int-3 of mouse, sequence of and epithelial cell transformation by
 Drosophila melanogaster...
 notch gene of, mouse mammary tumor gene int-3 as member of family of, sequence of and epithelial cell transformation by
 Protein sequences...
 of gene int-3 protein, of mouse mammary tumor
 Gene, animal, int-3...
 of mouse mammary tumor, sequence of and epithelial cell transformation by
 Deoxyribonucleic acids, repetitive...
 Saccharomyces cerevisiae cell cycle gene cdc-10 homolog, in gene int-3 of mouse mammary tumor
 CAS REGISTRY NUMBERS:
 146991-60-8 amino acid sequence of, complete
 139861-79-3 nucleotide sequence of

16/7/176 (Item 3 from file: 399)
 DIALOG(R) File 399:CA SEARCH(R)
 (c) 1996 American Chemical Society. All rts. reserv.

120130975 CA: 120(11)130975d DISSERTATION
 TAN-1, the human homolog of Drosophila "Notch", is involved in chromosomal translocations in human lymphoblastic neoplasma

AUTHOR(S): Ellisen, Leif William
LOCATION: Stanford Univ., Stanford, CA, USA
DATE: 1992 PAGES: 80 pp. CODEN: DABBBA LANGUAGE: English CITATION:
Diss. Abstr. Int. B 1993, 53(7), 3307 AVAIL: Univ. Microfilms Int., Order
No. DA9234046

SECTION:

CA214001 Mammalian Pathological Biochemistry

CA203XXX Biochemical Genetics

IDENTIFIERS: gene TAN1 chromosome translocation lymphoblastic leukemia

DESCRIPTORS:

Leukemia, T-cell acute lymphocytic...

gene TAN-1 translocation in, in human

Recombination, genetic, translocation...

of gene TAN-1, in human T-cell acute lymphoblastic leukemia

Gene, animal...

TAN-1, chromosomal translocation of, in human T-cell lymphoblastic
leukemia

Chromosome, human 7...

TAN-1 gene translocation from, in human T-cell lymphoblastic leukemia

16/7/141 (Item 28 from file: 155)
DIALOG(R)File 155:MEDLINE(R)
(c) format only 1996 Knight-Ridder Info. All rts. reserv.

06701790 89003790

Expression of the differentiation antigen F7D6 in tumorous tissues of *Drosophila*.

Bedian V; Jungklaus CE

Department of Biology, Clarkson University, Potsdam, New York 13676.

Dev Genet (UNITED STATES) 1987, 8 (3) p165-77, ISSN 0192-253X

Journal Code: DEG

Languages: ENGLISH

Document type: JOURNAL ARTICLE

The 63-kDa antigen recognized by the monoclonal antibody F7D6 is present in all *Drosophila* embryonic cells and disappears from most tissues as each one reaches its final, differentiated state. Larval tissues lose the antigen around the time of hatching, imaginal tissues lose it during metamorphosis, and germ cells lose it during gametogenesis (Bedian et al: Devel Biol 115:105-118, 1986). The nervous system and spontaneously contracting musculature of the gut and gonads are exceptions and remain antigen positive at all stages. The F7D6 antigen appears to be associated with dividing, undifferentiated cells and electrogenic cells. This prompted us to test tumors for antigen presence. We tested four different recessive mutants that give rise to four different types of tumorous transformation: the embryonic tumor Notch, several larval melanotic tumors, the imaginal disc tumor 1(2)gl, and three alleles of the ovarian tumor otu. In all cases, tumorous tissues in homozygotes contained the F7D6 antigen. The electrophoretic mobility of the antigen appeared to be unaltered in tumorous tissues compared to normal cells, but the antigen is expressed at higher levels. The antigen is found on the cytoplasmic surface of plasma membranes and appears to be a marker of undifferentiated normal and tumorous cells. Similarities and differences between the F7D6 antigen and *Drosophila* c-src protein are discussed.

16/7/139 (Item 26 from file: 155)
DIALOG(R)File 155:MEDLINE(R)
(c) format only 1996 Knight-Ridder Info. All rts. reserv.

06969982 89271982

[Adult T-cell leukemia (ATL)]

Takatsuki K

Kumamoto University Medical School.

Gan To Kagaku Ryoho (JAPAN) Apr 1989, 16 (4 Pt 2-1) p880-5, ISSN

0385-0684 Journal Code: 6T8

Languages: JAPANESE Summary Languages: ENGLISH

Document type: JOURNAL ARTICLE English Abstract

Adult T-cell leukemia (ATL) was first discovered and reported in Japan, where it has a high incidence in the southwest region. The retrovirus HTLV-I (human T-cell lymphotropic virus type I) is considered to be related to its etiology. In ATL endemic areas, HTLV-I carriers are found at a fairly high percentage even among healthy individuals. ATL shows diverse clinical features. It can be divided into 5 types (acute type, chronic type, smoldering type, crisis type, and lymphoma type). ATL cells originate from the CD4-positive subset of peripheral T cells; they show a characteristic notch in the nucleus and a lobulation tendency. ATL resists chemotherapy, and patients with acute and lymphoma types have quite a poor

prognosis. A definite diagnosis of ATL is made by documenting the presence of HTLV-I proviral DNA in the DNA of tumor cells. HTLV-I infection is caused by transmission of live lymphocytes via three routes (from mother to children, from males to females, and by transfusion). Familial occurrence of ATL is frequently seen. HTLV-I infection is also seen in other countries, but its incidence is highest in Japan. It is thus an urgent task for Japanese physicians to eliminate HTLV-I infection.

16/7/128 (Item 15 from file: 155)
DIALOG(R) File 155:MEDLINE(R)
(c) format only 1996 Knight-Ridder Info. All rts. reserv.

08122975 92260975
Cancer, chromosomes, and genes.
Nowell PC
Department of Pathology and Laboratory Medicine, University of Pennsylvania, School of Medicine, Philadelphia.
Lab Invest (UNITED STATES) Apr 1992, 66 (4) p407-17, ISSN 0023-6837
Journal Code: KZ4
Languages: ENGLISH
Document type: JOURNAL ARTICLE; REVIEW; REVIEW, ACADEMIC
(67 Refs.)

16/7/123 (Item 10 from file: 155)
DIALOG(R) File 155:MEDLINE(R)
(c) format only 1996 Knight-Ridder Info. All rts. reserv.

09041700 94356700
Analysis of phenotypic abnormalities and cell fate changes caused by dominant activated and dominant negative forms of the Notch receptor in Drosophila development.
Rebay I; Fortini ME; Artavanis-Tsakonas S
Howard Hughes Medical Institute, Department of Cell Biology, Yale University, New Haven, Connecticut 06536-0812.
C R Acad Sci III (FRANCE) Sep 1993, 316 (9) p1097-123, ISSN 0764-4469
Journal Code: CA1
Languages: ENGLISH, FRENCH
Document type: JOURNAL ARTICLE; REVIEW; REVIEW, TUTORIAL
The Notch gene of Drosophila plays an important role in cell fate specification throughout development. The Notch protein contains a large extracellular domain of 36 EGF-like repeats as well as 3 Notch/lin-12 repeats and an intracellular domain with 6 cdc10/ankyrin repeats, motifs which are highly conserved in several vertebrate Notch homologues [1-7]. In this review we summarize the results of two recent studies which attempt to establish structure-function relationships of the various domains of the Notch gene product [8, 9]. The functions of various structural domains of the Notch protein in vivo were investigated using a series of deletion mutants which have been ectopically expressed either under the hsp70 heat-shock promoter or under the sevenless eye-specific promoter. Truncation of the extracellular domain of Drosophila Notch produces an activated receptor as judged by its ability to cause phenotypes matching those of gain-of-function alleles or duplications of the Notch locus [8]. Equivalent truncations of vertebrate Notch-related proteins have been associated with malignant neoplasms and other developmental abnormalities [3, 6, 10, 11]. In contrast, dominant negative phenotypes result from overexpression of a protein lacking most intracellular sequences. These

results were extended by an analysis of activated Notch function at single-cell resolution in the *Drosophila* compound eye [9]. It was shown that while overexpression of full-length Notch in defined cell types has no apparent effects, overexpression of activated Notch in the same cells transiently blocks their proper cell-fate commitment, causing them to either adopt incorrect cell fates or to differentiate incompletely. Moreover, an activated Notch protein lacking the transmembrane domain is translocated to the nucleus, raising the possibility that Notch may participate directly in nuclear events. (68 Refs.)

16/7/116 (Item 3 from file: 155)
DIALOG(R)File 155:MEDLINE(R)
(c) format only 1996 Knight-Ridder Info. All rts. reserv.

09600265 96121865
Oncogene activation and oncogene cooperation in MMTV-induced mouse mammary cancer.
van Leeuwen F; Nusse R
Howard Hughes Medical Institute, Department of Developmental Biology, Beckman Center, Stanford University, Medical Center, CA 94305-5428, USA.
Semin Cancer Biol (UNITED STATES) Jun 1995, 6 (3) p127-33, ISSN 1044-579X Journal Code: A6Y
Languages: ENGLISH
Document type: JOURNAL ARTICLE; REVIEW; REVIEW, TUTORIAL
In this short review, we will give an overview of the various genes that can be activated by insertion of proviral DNA of the mouse mammary tumor virus during the formation of mammary cancer. These genes fall within three families: Wnt genes, FGF genes and Notch-related genes. We will summarize our current understanding of the roles of these genes in tumorigenesis and in normal development, and the mechanisms of action of their gene products. Finally, we will give some examples of cooperation between these genes in various biological settings. (91 Refs.)

16/7/113 (Item 6 from file: 144)
DIALOG(R)File 144:Pascal
(c) 1996 INIST/CNRS. All rts. reserv.

04166645 PASCAL No.: 75-0002761
DEVELOPMENTAL CAPACITIES OF BENIGN AND MALIGNANT NEOPLASMS OF DROSOPHILA
GATEFF E; SCHNEIDERMAN H A
CENT. PATHOBIOL., UNIV. CALIFORNIA, IRVINE, CALIF. 92664, USA
Journal: ARCH. (WILHELM ROUX) ENTWICKL.-MECH. ORG., 1974, 176 (1) 23-65
Availability: CNRS-3573
No. of Refs.: 4P.1/2
Document Type: P (SERIAL) ; A (ANALYTIC)
Country of Publication: FEDERAL REPUBLIC OF GERMANY
Language: ENGLISH Summary Language: GERMAN

16/7/112 (Item 5 from file: 144)
DIALOG(R)File 144:Pascal
(c) 1996 INIST/CNRS. All rts. reserv.

10278296 PASCAL No.: 92-0484208
Expression pattern of Motch, a mouse homolog of *Drosophila* notch, suggests an important role in early postimplantation mouse development

DEL AMO F F; SMITH D E; SWIATEK P J; GENDRON-MAGUIRE M; GREENSPAN R J;
MCMAHON A P; GRIDLEY T
Roche inst. molecular biology, dep. cell developmental biology, Nutley NJ
07110, USA

Journal: Development : (Cambridge), 1992, 115 (3) 737-744
ISSN: 0950-1991 Availability: INIST-7560; 354000020190410090

No. of Refs.: 1 p.1/2

Document Type: P (Serial) ; A (Analytic)

Country of Publication: United Kingdom

Language: English

The Notch gene of Drosophila encodes a large transmembrane protein involved in cell-cell interactions and cell fate decisions in the Drosophila embryo. To determine if a gene homologous to Drosophila Notch plays a role in early mouse development, we screened a mouse embryo cDNA library with probes from the Xenopus Notch homolog, Xotch. A partial cDNA clone encoding the mouse Notch homolog, which we have termed Motch, was used to analyze expression of the Motch gene. Motch transcripts were detected in a wide variety of adult tissues, which included derivatives of all three germ layers

16/7/110 (Item 3 from file: 144)
DIALOG(R)File 144:Pascal
(c) 1996 INIST/CNRS. All rts. reserv.

12471360 PASCAL No.: 96-0134151

Epithelial expression and chromosomal location of human TLE genes :
implications for notch signaling and neoplasia

YANLING LIU; GHASSAN DEHNI; PURCELL K J; SOKOLOW J; CARCANGIU M L;
ARTAVANIS-TSAKONAS S; STIFANI S

McGill univ., Montreal neurological inst., Montreal PQ H3A 2B4, Canada

Journal: Genomics : (San Diego, CA), 1996, 31 (1) 58-64

ISSN: 0888-7543 Availability: INIST-21389; 354000052959520080

No. of Refs.: 26 ref.

Document Type: P (Serial) ; A (Analytic)

Country of Publication: USA

Language: English

The TLE genes are the human homologues of Drosophila groucho, a member of the Notch signaling pathway. This pathway controls a number of different cell-fate choices in invertebrates and vertebrates. We are interested in investigating the functions of the TLE gene family during epithelial determination and carcinogenesis. We show that expression of individual TLE genes correlates with immature epithelial cells that are progressing toward their terminally differentiated state, suggesting a role during epithelial differentiation. In both normal tissues and conditions resulting from incorrect or incomplete maturation events, such as metaplastic and neoplastic transformations, TLE expression is elevated and coincides with Notch expression, implicating these molecules in the maintenance of the undifferentiated state in epithelial cells. We also show that TLE1 and TLE2 are organized in a tandem array at chromosomal location 19p13.3, while TLE3 maps to 15q22.

?t s16/7/2,3,6,9,10,11,12,14,15,17,23,25,40

>>>Format 7 is not valid in file 143

16/7/2 (Item 2 from file: 5)
DIALOG(R)File 5:BIOSIS PREVIEWS(R)
(c) 1996 BIOSIS. All rts. reserv.

13041777 BIOSIS Number: 99041777

Exclusive development of T cell neoplasms in mice transplanted with bone marrow expressing activated Notch alleles

Pear W S; Aster J C; Scott M L; Hassler R P; Soffer B; Sklar J; Baltimore D

Massachusetts Inst. Technol., Room 68-380, 77 Massachusetts Ave., Cambridge, MA 02139-4307, USA

Journal of Experimental Medicine 183 (5). 1996. 2283-2291.

Full Journal Title: Journal of Experimental Medicine

ISSN: 0022-1007

Language: ENGLISH

Print Number: Biological Abstracts Vol. 102 Iss. 002 Ref. 023950

Notch is a highly conserved transmembrane protein that is involved in cell fate decisions and is found in organisms ranging from *Drosophila* to humans. A human homologue of Notch, TAN1, was initially identified at the chromosomal breakpoint of a subset of T-cell lymphoblastic leukemias/lymphomas containing a t(7;9) chromosomal translocation; however, its role in oncogenesis has been unclear. Using a bone marrow reconstitution assay with cells containing retrovirally transduced TAN1 alleles, we analyzed the oncogenic potential of both nuclear and extranuclear forms of truncated TAN1 in hematopoietic cells. Although the Moloney leukemia virus long terminal repeat drives expression in most hematopoietic cell types, retroviruses encoding either form of the TAN1 protein induced clonal leukemias of exclusively immature T cell phenotypes in approx 50% of transplanted animals. All tumors overexpressed truncated TAN1 of the size and subcellular localization predicted from the structure of the gene. These results show that TAN1 is an oncoprotein and suggest that truncation and overexpression are important determinants of transforming activity. Moreover, the murine tumors caused by TAN1 in the bone marrow transplant model are very similar to the TAN1-associated human tumors and suggest that TAN1 may be specifically oncogenic for T cells.

16/7/3 (Item 3 from file: 5)

DIALOG(R)File 5:BIOSIS PREVIEWS(R)

(c) 1996 BIOSIS. All rts. reserv.

13034202 BIOSIS Number: 99034202

T cell leukemia-associated human notch-translocation-associated notch homologue has I-kappa-B-like activity and physically interacts with nuclear factor-kappa-B proteins in T cells

Guan E; Wang J; Laborda J; Norcross M; Baeuerle P A; Hoffman T

Lab. Cell Biol., Div. Monoclonal Antibodies, Cent. Biologics Eval. Res., Federal Drug Administration, Building 29B, Room 4G03, HFM558, 8800 Rockville Pike, Bethesda, MD 20892, USA

Journal of Experimental Medicine 183 (5). 1996. 2025-2032.

Full Journal Title: Journal of Experimental Medicine

ISSN: 0022-1007

Language: ENGLISH

Print Number: Biological Abstracts Vol. 102 Iss. 002 Ref. 016375

Translocation-associated Notch homologue (TAN-1), a gene originally cloned from the translocation breakpoint of a human T cell leukemia carrying a 9:7(q34.3) translocation, encodes a protein belonging to the Notch/Lin-12/Glp-1 receptor family. These receptors mediate the specification of numerous cell fates during development in invertebrates and vertebrates. The intracellular portion of Notch/TAN-1 contains six ankyrin repeats that are similar to those found in cytoplasmic I-kappa-B

proteins. I-kappa-B proteins are specific inhibitors of nuclear factor (NF)-kappa-B/Rel transcription factors. Here we show that TAN-1 has functional properties of an I-kappa-B-like regulator with specificity for the NF-kappa-B p50 subunit. A recombinant polypeptide corresponding to the cytoplasmic portion of TAN-1 (TAN-1-C) specifically inhibited the DNA binding of p50-containing NF-kappa-B complexes. When overexpressed in an appropriate cell line, TAN-1-C prevented kappa-B-dependent transactivation in transient reporter gene assays in a fashion similar to the structurally related protein, Bcl-3. TAN-1-C could activate kappa-B-dependent gene expression by attenuating the inhibitory effect of an excess of p50 homodimers. Immunoprecipitation experiments showed that the TAN-1 from a T cell line is associated with NF-kappa-B containing p50 and p65 subunits. These observations indicate that TAN-1-C may directly engage NF-kappa-B transcription factors and modulate nuclear gene expression.

16/7/6 (Item 6 from file: 5)
DIALOG(R)File 5:BIOSIS PREVIEWS(R)
(c) 1996 BIOSIS. All rts. reserv.

12098006 BIOSIS Number: 98698006

Epithelial expression and chromosomal location of human TLE genes:
Implications for notch signaling and neoplasia

Liu Y; Dehni G; Purcell K J; Sokolow J; Carcangiu M L; Artavanis-Tsakonas S; Stifani S

Montreal Neurol. Inst., 3801 University St., Montreal, PQ H3A 2B4, Canada
Genomics 31 (1). 1996. 58-64.

Full Journal Title: Genomics

ISSN: 0888-7543

Language: ENGLISH

Print Number: Biological Abstracts Vol. 101 Iss. 007 Ref. 098287

The TLE genes are the human homologues of Drosophila groucho, a member of the Notch signaling pathway. This pathway controls a number of different gene family during epithelial determination and carcinogenesis. We show that expression of individual TLE genes correlates with immature epithelial cells that are progressing toward their terminally differentiated state, suggesting a role during epithelial differentiation. In both normal tissues and conditions resulting from incorrect or incomplete maturation events, such as metaplastic and neoplastic transformations, TLE expression is elevated and coincides with Notch expression, implicating these molecules in the maintenance of the undifferentiated state in epithelial cells. We also show that TLE1 and TLE2 are organized in a tandem array at chromosomal location 19p13.3, while TLE3 maps to 15q22.

16/7/9 (Item 9 from file: 5)
DIALOG(R)File 5:BIOSIS PREVIEWS(R)
(c) 1996 BIOSIS. All rts. reserv.

11799009 BIOSIS Number: 98399009

TAN-1, human homolog of Drosophila Notch, encodes a IKB-like activity
Guan E; Wang J H; Laborda J; Lizzio E; Norcross M; Baeuerle P A; Hoffman T

Cent. Biol. Eval. Res., US FDA, Freiburg, Germany

0 (0). 1995. 693.

Full Journal Title: 9TH INTERNATIONAL CONGRESS OF IMMUNOLOGY. The 9th International Congress of Immunology; Meeting Sponsored by the American Association of Immunologists and the International Union of Immunological

Societies, San Francisco, California, USA, July 23-29, 1995. ix+742p. 9th International Congress of Immunology: San Francisco, California, USA.

ISSN: *****

Language: ENGLISH

Print Number: Biological Abstracts/RRM Vol. 047 Iss. 009 Ref. 161412

16/7/10 (Item 10 from file: 5)
DIALOG(R)File 5:BIOSIS PREVIEWS(R)
(c) 1996 BIOSIS. All rts. reserv.

11789761 BIOSIS Number: 98389761

Functional analysis of the TAN-1 gene, a human homolog of Drosophila notch

Aster J; Pear W; Hasserjian R; Erba H; Davi F; Luo B; Scott M; Baltimore D; Sklar J

Div. Molecular Oncol., Dep. Pathol., Brigham and Women's Hosp., Harv. Med. Sch., Boston, MA 02115, USA

0 (0). 1994. 125-136.

Full Journal Title: COLD SPRING HARBOR LABORATORY. Cold Spring Harbor Symposia on Quantitative Biology, Vol. 59. The molecular genetics of cancer; 59th Symposium on Quantitative Biology, Cold Spring Harbor, New York, USA, June 1-8, 1994. xxvii+739p. Cold Spring Harbor Laboratory Press: Plainview, New York, USA. ISBN 0-87969-068-2(paper); ISBN 0-87969-067-4(cloth).

ISSN: 0091-7451

Language: ENGLISH

Print Number: Biological Abstracts/RRM Vol. 047 Iss. 009 Ref. 152164

16/7/11 (Item 11 from file: 5)
DIALOG(R)File 5:BIOSIS PREVIEWS(R)
(c) 1996 BIOSIS. All rts. reserv.

11778798 BIOSIS Number: 98378798

Alterations in Notch signaling in neoplastic lesions of the human cervix
Zagouras P; Stifani S; Blaumueller C M; Carcangiu M L; Artavanis-Tsakonas S

Dep. Biol., Yale Univ., New Haven, CT 06536, USA

Proceedings of the National Academy of Sciences of the United States of America 92 (14). 1995. 6414-6418.

Full Journal Title: Proceedings of the National Academy of Sciences of the United States of America

ISSN: 0027-8424

Language: ENGLISH

Print Number: Biological Abstracts Vol. 100 Iss. 005 Ref. 070636

The development of cancer is a cellular process that reflects and is partly driven by alterations in cell determination. Mutations in various molecules responsible for cell determination have been identified as being oncogenic, but little is known about the involvement of normal cell fate-determining mechanisms in the oncogenic process. The Notch pathway defines an evolutionarily conserved, general cell interaction mechanism that controls fundamental aspects of cell determination during vertebrate and invertebrate development. We have explored the involvement of the human Notch pathway in human cervical tissues, which define a cellular environment where cell fate changes take place and where neoplastic conditions have been well characterized. Our evidence suggests that Notch expression is associated with cell populations that are undergoing cell

fate changes and that Notch activity can be used to monitor cell fate abnormalities in cervical as well as other epithelial neoplasias.

16/7/12 (Item 12 from file: 5)
DIALOG(R)File 5:BIOSIS PREVIEWS(R)
(c) 1996 BIOSIS. All rts. reserv.

11572298 BIOSIS Number: 98172298

Drosophila in cancer research: The first fifty tumor suppressor genes

Watson K L; Justice R W; Bryant P J

Dep. Molecular Cellular Biol., Harvard Univ., Cambridge, MA 02138, USA

Journal of Cell Science 0 (SUPPL. 18). 1994. 19-33.

Full Journal Title: Journal of Cell Science

ISSN: 0021-9533

Language: ENGLISH

Print Number: Biological Abstracts Vol. 099 Iss. 008 Ref. 112601

Print Number: Biological Abstracts/RRM Vol. 047 Iss. 005 Ref. 069789

In Drosophila, over 50 genes have been identified in which loss-of-function mutations lead to excess cell proliferation in the embryo, in the central nervous system, imaginal discs or hematopoietic organs of the larva, or in the adult gonads. Twenty-two of these genes have been cloned and characterized at the molecular level, and nine of them show clear homology to mammalian genes. Most of these mammalian genes had not been previously implicated in cell proliferation control. Overgrowth in some of the mutants involves conversion to a cell type that, in normal development, shows more cell proliferation than the original cell type. Thus the neurogenic mutants, including Notch, show conversion of epidermal cells to neuroblasts, leading to the 'neurogenic' phenotype of excess nervous tissue. The ovarian tumor mutants show conversion of the female germ line to a cell type resembling the male germ line, which undergoes more proliferation than the female germ line. Mutations of the fat locus cause hyperplastic overgrowth of imaginal discs, in which the epithelial structure is largely intact. The predicted fat protein product is a giant relative of cadherins, supporting indications from human cancer that cadherins play an important role in tumor suppression. Mutations in the lethal(2)giant larvae and lethal(1)discs large genes cause neoplastic overgrowth of imaginal discs as well as the larval brain. The dig gene encodes a membrane-associated guanylate kinase homolog that is localized at septate junctions between epithelial cells. This protein is a member of a family of homologs that also includes two proteins found at mammalian tight junctions (ZO-1 and ZO-2) and a protein found at mammalian synaptic junctions (PSD-95/SAP90). Genes in which mutations cause blood cell overproduction include aberrant immune response-8, which encodes the RpS6 ribosomal protein and hopscotch, which encodes a putative non-receptor protein tyrosine kinase. The gene products identified by ovarian tumor mutants do not show clear amino acid sequence homology to known proteins. Drosophila provides an opportunity to rapidly identify and characterize tumor suppressor genes, many of which have mammalian homologs that might also be involved in cell proliferation control and tumor suppression.

16/7/14 (Item 14 from file: 5)
DIALOG(R)File 5:BIOSIS PREVIEWS(R)
(c) 1996 BIOSIS. All rts. reserv.

11492147 BIOSIS Number: 98092147

The human NOTCH1, 2, and 3 genes are located at chromosome positions

9q34, 1p13-p11, and 19p13.2-p13.1 in regions of neoplasia-associated translocation

Larsson C; Lardelli M; White I; Lendahl U

Dep. Cell Mol. Biol., Med. Nobel Inst., Karolinska Hosp., S-171 77 Stockholm, Sweden

Genomics 24 (2). 1994. 253-258.

Full Journal Title: Genomics

ISSN: 0888-7543

Language: ENGLISH

Print Number: Biological Abstracts Vol. 099 Iss. 005 Ref. 062557

In *Drosophila* the Notch gene controls differentiation to various cell fates in many tissues. Three mammalian Notch homologs have recently been identified: Notch 1, 2, and 3. All three homologs are very highly conserved relative to the *Drosophila* Notch gene, which suggests that they are important for cell differentiation in mammals. This notion is supported by the previous finding of a truncated, translocated form of the human NOTCH1 gene (formerly TAN1) in three cases of leukemia. Given this genetic link between NOTCH1 and tumor formation, it is of interest to establish the chromosomal positions of the other two homologs. We report the identification of cosmid clones for the human NOTCH1, 2, and 3 genes. These clones were used as probes in fluorescence in situ hybridization to human metaphase chromosomes, and the results, combined with data from somatic cell hybrid panels, show that the NOTCH2 and 3 genes are located at positions 1p13-p11 and 19p13.2-p13.1, respectively, which are regions of neoplasia-associated translocation.

16/7/15 (Item 15 from file: 5)

DIALOG(R)File 5:BIOSIS PREVIEWS(R)

(c) 1996 BIOSIS. All rts. reserv.

11292950 BIOSIS Number: 97492950

An activated Notch suppresses neurogenesis and myogenesis but not gliogenesis in mammalian cells

Nye J S; Kopan R; Axel R

Dep. Neurol., Inst. Cancer Res., Coll. Physicans Surg., Columbia Univ., New York, NY 10032, USA

Development (Cambridge) 120 (9). 1994. 2421-2430.

Full Journal Title: Development (Cambridge)

ISSN: 0950-1991

Language: ENGLISH

Print Number: Biological Abstracts Vol. 098 Iss. 010 Ref. 128689

P19 cells, a mouse embryonal carcinoma line, can be induced to differentiate into neurons. After induction, however, only a small subpopulation of cells develop as neurons, suggesting that equipotent cells adopt different cell fates. In invertebrate systems, the lin-12-Notch family of genes is thought to control the choice of cell fate. We have therefore asked whether activation of murine Notch (mNotch) regulates neuronal differentiation in P19 cells. We demonstrate that a dominant gain-of-function mutant of mNotch suppresses neurogenesis, as well as myogenesis in P19 cells. Overexpression of the full-length mNotch protein also suppresses neurogenesis. In contrast, the differentiation of glia is not affected by an activated mNotch homologue. These data indicate that mNotch may play a central role in the choice of cell fate in differentiating cells in culture and suggests that mNotch may play a similar role in the choice of fate in the developing mammalian embryo.

16/7/17 (Item 17 from file: 5)
DIALOG(R)File 5:BIOSIS PREVIEWS(R)
(c) 1996 BIOSIS. All rts. reserv.

11038694 BIOSIS Number: 97238694

A human homologue of the Drosophila developmental gene, Notch, is expressed in CD34+ hematopoietic precursors

Milner L A; Kopan R; Martin D I K; Bernstein I D

Fred Hutchinson Cancer Res. Cent., 1100 Fairview Ave. N, Seattle WA 98109, USA

Blood 83 (8). 1994. 2057-2062.

Full Journal Title: Blood

ISSN: 0006-4971

Language: ENGLISH

Print Number: Biological Abstracts Vol. 097 Iss. 011 Ref. 155807

Members of the Notch gene family have been shown to mediate cell-fate decisions by multipotent precursors in a number of different systems. To determine whether members of this family might play a similar role in hematopoiesis, we asked if homologues of the Notch gene are expressed in human hematopoietic precursors. Using degenerate oligonucleotides corresponding to conserved amino acid sequences in known Notch homologues as primers for the polymerase chain reaction (PCR), we demonstrated that at least one Notch homologue is expressed in human bone marrow CD34+ cells, a population enriched for hematopoietic precursors. Cloning and sequencing of the PCR products identified this Notch homologue as TAN-1, a member of the Notch family previously cloned from a T-cell leukemia with a translocation involving this gene. Subsequent evaluation of bone marrow hematopoietic cells for TAN-1 expression using a reverse transcription-PCR assay confirmed the expression of TAN-1 in CD34+ hematopoietic precursors, including the immature subset that lacks expression of lineage-associated antigens (CD34+lin-). These findings, together with the known role of Notch homologues in other systems, suggest that members of the Notch family, including TAN-1, may be involved in mediating cell-fate decisions during hemotopoiesis.

16/7/23 (Item 23 from file: 5)
DIALOG(R)File 5:BIOSIS PREVIEWS(R)
(c) 1996 BIOSIS. All rts. reserv.

9139819 BIOSIS Number: 93124819

EXPRESSION OF AN ACTIVATED NOTCH-RELATED INT-3 TRANSGENE INTERFERES WITH CELL DIFFERENTIATION AND INDUCES NEOPLASTIC TRANSFORMATION IN MAMMARY AND SALIVARY GLANDS

JHAPPAN C; GALLAHAN D; STAHL C; CHU E; SMITH G H; MERLINO G; CALLAHAN R
LAB. MOLECULAR BIOL., NATL. CANCER INST., NATL. INST. HEALTH. BETHESDA, MD. 20892, USA.

GENES DEV 6 (3). 1992. 345-355. CODEN: GEDEE

Full Journal Title: Genes & Development

Language: ENGLISH

Expression of the int-3 locus is activated in mouse mammary tumors as a consequence of insertional mutagenesis by the mouse mammary tumor virus (MMTV). Integration of the MMTV provirus into the into-3 locus promotes the transcription and translation of flanking cellular int-3 sequences sharing significant homology with the intracellular domain of the nuerogenic Notch gene of Drosophila, and with the yeast cell cycle regulatory genes cdc10 and SW16. To determine the in vivo consequences of activated int-3 expression, transgenic mice were generated harboring a genomic tumor DNA

fragment consisting of the MMTV LTR and the flanking cellular int-3 sequences. All six int-3 founder transgenic mice and the progeny of one established line exhibited similar dramatic phenotypic abnormalities in tissues in which the transgene was expressed. Focal and often multiple poorly differentiated mammary and salivary adenocarcinomas appeared in the majority of transgenic mice between 2 and 7 months of age. Significantly, mammary glands were arrested in development and were lactation deficient in all female int-3 mice. The salivary glands, glands of the nasal mucosa and maxillary sinus, the extraorbital lacrimal glands, and the Harderian glands of juvenile and adult transgenic mice all contained proliferating immature ductule cells and were incompletely differentiated. In addition, all male int-3 transgenic mice were sterile, apparently the result of severe hyperplasia of the epididymis. These findings demonstrate in vivo that expression of the activated Notch-related int-3 gene causes deregulation of normal developmental controls and hyperproliferation of glandular epithelia.

16/7/25 (Item 25 from file: 5)
DIALOG(R)File 5:BIOSIS PREVIEWS(R)
(c) 1996 BIOSIS. All rts. reserv.

8646875 BIOSIS Number: 92111875
TAN-1 THE HUMAN HOMOLOG OF THE DROSOPHILA NOTCH GENE IS BROKEN BY
CHROMOSOMAL TRANSLOCATIONS IN T LYMPHOBLASTIC NEOPLASMS
ELLISEN L W; BIRD J; WEST D C; SORENG A L; REYNOLDS T C; SMITH S D; SKLAR
J

STANFORD UNIV. SCH. MED., STANFORD, CALIF. 94305.

CELL 66 (4). 1991. 649-662. CODEN: CELLB

Full Journal Title: Cell

Language: ENGLISH

Previously we described joining of DNA in the .beta. T cell receptor gene to DNA of an uncharacterized locus in a t(7;9)(q34;q34.3) chromosomal translocation from a case of human lymphoblastic leukemia (T-ALL). We now show that the locus on chromosome 9 contains a gene highly homologous to the Drosophila gene Notch. Transcripts of the human gene, for which we propose the name TAN-1, and its murine counterpart are present in many normal human fetal and adult mouse tissues, but are most abundant in lymphoid tissues. In t(7;9)(q34;q34.3) translocations from three cases of T-ALL, the breakpoints occur within 100 bp of an intron in TAN-1, resulting in truncation of TAN-1 transcripts. These observations suggest that TAN-1 may be important for normal lymphocyte function and that alteration of TAN-1 may play a role in the pathogenesis of some T cell neoplasms.

16/7/40 (Item 1 from file: 73)
DIALOG(R)File 73:EMBASE
(c) 1996 Elsevier Science B.V. All rts. reserv.

9978677 EMBASE No: 96166337
T cell leukemia-associated human Notch/translocation-associated Notch homologue has IkappaB-like activity and physically interacts with nuclear factor-kappaB proteins in T cells

Guan E.; Wang J.; Laborda J.; Norcross M.; Baeuerle P.A.; Hoffman T.

Division of Monoclonal Antibodies, Federal Drug Administration, Ctr. for Biologics Evaluation/Res., 8800 Rockville Pike, Bethesda, MD 20892 USA

Journal of Experimental Medicine (USA) , 1996, 183/5 (2025-2032) CODEN: JEMEA
ISSN: 0022-1007

LANGUAGES: English SUMMARY LANGUAGES: English

Translocation-associated Notch homologue (TAN-1), a gene originally cloned from the translocation breakpoint of a human T cell leukemia carrying a 9:7(q34.3) translocation, encodes a protein belonging to the Notch/Lin-12/Glp-1 receptor family. These receptors mediate the specification of numerous cell fates during development in invertebrates and vertebrates. The intracellular portion of Notch/TAN-1 contains six ankyrin repeats that are similar to those found in cytoplasmic IkappaB proteins. IkappaB proteins are specific inhibitors of nuclear factor (NF)-kappaB/Rel transcription factors. Here we show that TAN-1 has functional properties of an IkappaB-like regulator with specificity for the NF-kappaB p50 subunit. A recombinant polypeptide corresponding to the cytoplasmic portion of TAN-1 (TAN-1(C)) specifically inhibited the DNA binding of p50-containing NF-kappaB complexes. When overexpressed in an appropriate cell line, TAN-1(C) prevented kappaB-dependent transactivation in transient reporter gene assays in a fashion similar to the structurally related protein, Bcl-3. TAN-1(C) could activate kappaB-dependent gene expression by attenuating the inhibitory effect of an excess of p50 homodimers. Immunoprecipitation experiments showed that the TAN-1 from a T cell line is associated with NF-kappaB containing p50 and p65 subunits. These observations indicate that TAN-1(C) may directly engage NF-kappaB transcription factors and modulate nuclear gene expression.

?LOGOFF